

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 4, 2002, 16:17:30 ; Search time 71.77 Seconds  
(without alignments)  
118.689 Million cell updates/sec

Title: US-09-052-089a-3  
Perfect score: 1066  
Sequence: 1 RTIINKLFEDLAQEEENVLD.....DLQASADKEIMSLKKLTMLQ 220

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	170.5	16.0	1957	YD86_SCHPO	Q10411 schizosacch
2	168	15.8	962	VDP_HUMAN	O60763 homo sapien
3	164.5	15.4	2116	MY52_DICDI	P08799 dictyosteli
4	163.5	15.3	1744	TANA_XENLA	Q01550 xenopus lae
5	161	15.1	1325	G160_MOUSE	P55937 mus musculu
6	160.5	15.1	1938	MY5_AEOIR	P41733 aequipecten
7	159	14.9	1790	US01_YEAST	P25386 saccharomyc
8	159	14.9	1938	MYHD_HUMAN	Q9ukx3 homo sapien
9	158.5	14.9	1935	MYH7_RAT	P02564 ratus norv
10	157	14.7	1935	MYSS_CYPCA	Q90339 cyprinus ca
11	155.5	14.6	941	VDP_MOUSE	Q92140 mus musculu
12	155.5	14.6	1203	XCPE_XENLA	P80533 xenopus lae
13	155	14.5	1938	MYH4_RABIT	Q28641 oryctolagus
14	155	14.5	1940	MYH3_RAT	P12847 ratus norv
15	155	14.5	3210	CENF_HUMAN	P49454 homo sapien
16	154.5	14.5	959	VDP_RAT	P15442 ratus norv
17	154.5	14.5	1509	MY5N_ACACA	P05659 acanthamoeb
18	154	14.4	1941	MYH2_HUMAN	Q9ukx2 homo sapien
19	152.5	14.3	579	G160_HUMAN	Q08378 homo sapien
20	152.5	14.3	1935	MYH7_HUMAN	P12883 homo sapien
21	152	14.3	2411	MYSA_DROME	P05661 drosophila
22	151	14.2	722	MEP1_TOBAC	Q9M734 nicotiana t
23	150	14.1	1969	MYSA_CAEEL	P12844 caenorhabdi
24	149	14.0	548	KUCR_MOUSE	P70194 mus musculu
25	149	14.0	919	RA50_AERPE	Q9Y121 aeropyrum p
26	149	14.0	978	RA50_AQUAE	O67124 aquilex aeo
27	149	14.0	1939	MYH1_HUMAN	P12882 homo sapien
28	148.5	13.9	886	RA50_ARCFU	Q29230 archaeoglob
29	148.5	13.9	963	KINH_HUMAN	P31376 homo sapien
30	148.5	13.9	1934	MYH7_MESAN	P33540 mesocricetu
31	148.5	13.9	1938	MYSS_CHICK	P13538 gallus galli
32	148.5	13.9	1959	MYH9_CHICK	P14105 gallus galli
33	148.5	13.9	1966	MY5B_CAEEL	P02566 caenorhabdi

34	147.5	13.8	2017	1	MY5N_DROME	Q99323 drosophila
35	147.5	13.8	2653	1	CENE_HUMAN	O02224 homo sapien
36	147	13.8	1453	1	Y373_BOVIN	O91u23 bos taurus
37	146.5	13.7	895	1	RA50_TREVO	P58302 thermoplasma
38	146	13.7	1084	1	MYSS_RABIT	P02562 oryctolagus
39	146	13.7	1290	1	XCPC_XENLA	P50532 xenopus lae
40	146	13.7	1940	1	MYH3_HUMAN	P11055 homo sapien
41	145.5	13.6	388	1	PAM_STRPY	P49054 streptococc
42	145.5	13.6	1937	1	MYH8_HUMAN	P13535 homo sapien
43	145.5	13.6	1960	1	MYH9_HUMAN	P35379 homo sapien
44	145	13.6	407	1	M21_STRPY	P50468 streptococc
45	144.5	13.6	961	1	VDP_BOVIN	P41541 bos taurus

## ALIGNMENTS

RESULT	ID	STANDARD	PRT
1	YD86_SCHPO	STANDARD	PRT: 1957 AA.
YD86_SCHPO	YD86_SCHPO		
Q10411:			
AC	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	Hypothetical 222.8 kDa protein C1F3.06C in chromosome 1.		
GN	SPAC1F3.06C.		
OS	Schizosaccharomyces pombe (fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;		
OC	Schizosaccharomycetes.		
ON	NCBI_TaxID=4896;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=972;		
RA	Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;		
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).		
CC	-----		
DR	EMBL; Z70690; CA94624.1; -		
KW	Hypothetical protein.		
SO	SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;		

Query Match	16.0%; Score 170.5; DB 1; Length 1957;
Best Local Similarity	22.6%; Pred. No. 0.027;
Matches	58; Conservative 53; Mismatches 92; Indels 43; Gaps 6;
QY	15 EENVLDREFLEKNEIDNVRAOLSKDEKRSQVITDRLTLEBRNATVSLQALGKAE 74
DB	1386 EDNQLATKAKLNQDHLNQLRKEKEDVLAKEKESLISLEESLSNQROKESLDAKNLE 1445
QY	75 -MLCSTLKKQMYLEQODDETKQAQEEAGRLRSKMTN-----EQIEEL--LQSLP 123
DB	1446 HMLDTSKRNLSLMEKIESINSLSDDKSFELASVEKLGALQKLHSELSIMENIKSOLQ 1505
QY	124 EEEEMIRMGQSGAVEDLAVYCVSLKKEYEN-----LKEARK 161
DB	1506 EAEKETO--VDESTIOELDEHETASKNNYSGKLNDKDSIIRDSENIEQLNNLLAEKRS 1562
QY	162 ASGEVADKLRLDLPSSRSKL-----QTVSELDQAKLELKSQKDLQASADKEIMSLK 214
DB	1563 AVKRILSTKESEIILQFNRLADLEVHKSQVSELSGRSLKLAISTTEELQLANERLSLT 1622
QY	215 KLTMLQ 220
DB	1623 RMLDLQ 1628

```

RESULT 2
VDP_HUMAN STANDARD: PRT: 962 AA.
ID VDP_HUMAN 060763;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE General vesicular transport factor p115 (Transcytosis associated
  protein) (TAP) (Vesicle docking protein).
GN VDP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION SITE SER-942.
RA MEDLINE=98148093; PubMed=9478999;
RX Sonda M., Misumi Y., Yano A., Takami N., Ikehara Y.;
RT "Phosphorylation of the vesicle docking protein p115 regulates its
  association with the Golgi membrane.";
RL J. Biol. Chem. 273:5385-5388(1998).
CC -I- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR
  INTERCISTERNAL TRANSPORT IN THE GOLGI STACK. IT IS REQUIRED FOR
  TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO
  THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY
  INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR
  AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES
  BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.
CC -I- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-
  COIL). AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
CC -I- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER:
  PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.
CC PHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE.
CC PHOSPHORYLATION PROMOTES DISSOCIATION.
CC -I- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D86326; BAA25300.1; -.
DR MIM: 603344; -.
DR InterPro: IPR000225; Armadillo.
DR PROSITE: PS50176; ARM_REPEAT; UNKNOWN_1.
KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil;
  KX Phosphorylation.
FT DOMAIN 1 637 GLOBULAR HEAD.
FT DOMAIN 638 930 COILED COIL (POTENTIAL).
FT DOMAIN 935 962 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 942 942 PHOSPHORYLATION.
FT MUTAGEN 942 942 S->A: LOSS OF PHOSPHORYLATION.
SQ SEQUENCE 962 AA: 107906 MW: 2E748F2C1BC2B942 CRC64;

```

```

Query Match 15.8%; Score 168; DB 1; Length 962;
Best Local Similarity 20.2%; Pred. No. 0.018;
Matches 65; Conservative 62; Mismatches 85; Indels 110; Gaps 7;

```

```

QY 3 IINKLFLPADBEENVLDREFLNELDNVRAQLSOKDEKRDSDVITLRLDTLEERNAT 62
DB 608 IFDEEFKFLVELEGVYIKAYKSEKKEBEVKKLEQHDN--IVHYKMKITREQLDQ 665
QY 63 VVSLQOALG---KAEMCSTLKKOMKYLEGOODE-----TKQAQ-- 98
DB 666 LEELRQOVSTLKQNEQLQTAQTQVOVSQIQHKDQYNLKLQLGKDNQHGSSYSEGAOMN 725

```

```

QY 99 ----EAGRLRSKMTMEQIELLOSLPEVEEMIRMGVQSA----- 138
DB 726 GIQPEIRGLREIREIELKRNQELLOSLTEKDSMIENKSSQTSQTNQSSAIVAROSE 785
QY 139 -VEQLAVYCVSLK----- 150
DB 786 QVAELKQELATLKSQNSQSVETKLOTEKQELLOKTEAFKSVQGETTIAFKTTD 845
QY 151 -----KEYENLKEARKASGEVADLRKDLFSSRSKLQTVSELDQAKLELKSQAKD 201
DB 846 VEGRLSALLQETKEKLEKNEIKALSEERTAIKEQLDSSNSTIAILOTREKOLELEITDSKE 905
QY 202 -----LQSADEKTEISLKKRL 216
DB 906 QDDLVLVLADDOQKITLSLKNKL 927

```

```

RESULT 3
MYS2_DICD1 STANDARD: PRT: 2116 AA.
ID MYS2_DICD1
AC P08799;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin II heavy chain, non muscle.
GN MHCA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87092266; PubMed=3540939;
RA Warrick H.M., de Iozanne A., Leitwand L.A., Spudis J.A.;
RT "Conserved protein domain in a myosin heavy chain gene from
  Dictyostelium discoideum.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN [2]
RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC STRAIN=AX2;
RX MEDLINE=90353583; PubMed=2387408;
RA Lucke-Vietmeier D., Schleicher M., Grabatin B., Wipplier J.,
  Gerisch G.;
RT "Replacement of threonine residues by serine and alanine in a
  phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL FEBS Lett. 269:239-243(1990).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE=88112226; PubMed=2828113;
RA Magle G., Noegel A., Scheel J., Gerisch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
  Dictyostelium myosin heavy chain.";
RL FEBS Lett. 227:71-75(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
  Rayment I.;
RT "X-ray structures of the myosin motor domain of Dictyostelium
  discoideum complexed with MgADP. Bepx and MgADP.ALf4-";
RL Biochemistry 34:8960-8972(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
  truncated head of Dictyostelium discoideum myosin to 2.7-A
  resolution.";
RL Biochemistry 34:8973-8981(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the

```

RT Dictyostelium discoidium myosin motor domain to 1.9-A resolution." ;  
 RL Biochemistry 35:5404-5417(1996).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
 RX MEDLINE=97452580; PubMed=9305951;  
 RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.,  
 RT "X-ray structures of the MgAD, MgATPgammaS, and MgAMPN complexes  
 of the Dictyostelium discoidium myosin motor domain." ;  
 RL J. Mol. Biol. 274:394-407(1997).  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE  
 CC ACTIVITY THAT IS ACTIVATED BY ACTIN.  
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
 CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI  
 CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
 CC (MLC-2').  
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL  
 CC CORTEX.  
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER  
 CC SPLIT INTO 2 GLOBULAR SUBDOMAINS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PPM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES  
 CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE  
 CC ACTIVITY. PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1  
 CC POSITION (668).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, M14628; AAA33227.1; -.  
 DR PIR, A26655; A26655.  
 DR PIR, S00250; S00250.  
 DR PDB, 1MMA: 03-DEC-97.  
 DR PDB, 1MMD: 17-AUG-96.  
 DR PDB, 1MMG: 03-DEC-97.  
 DR PDB, 1MMN: 03-DEC-97.  
 DR PDB, 1MND: 17-AUG-96.  
 DR PDB, 1MNE: 17-AUG-96.  
 DR PDB, 1VOM: 23-DEC-96.  
 DR PDB, 1LVK: 28-JAN-98.  
 DR Dictydb, DD01008; mhca.  
 DR InterPro, IPR000048; IQ.  
 DR InterPro, IPR004009; Myosin\_N.  
 DR InterPro, IPR001609; myosin\_head.  
 DR Pfam, PF00612; IQ; 2.  
 DR Pfam, PF00063; myosin\_head; 1.  
 DR Pfam, PF02736; Myosin\_N; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR ProDom, PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 KW Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;  
 KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.

```

FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
FT DOMAIN 762 791 IQ.
FT DOMAIN 817 2116 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 638 660 ACTIN-BINDING.
FT DOMAIN 738 752 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
FT MOD_RES 678 678 ALKYLATION (SH-1).
FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BB1EE56A1 CRC64;

Query Match 15.4%, Score 164.5; DB 1; Length 2116;
Best Local Similarity 25.4%; Pred. No. 0.063;
Matches 68; Conservative 56; Mismatches 87; Indels 57; Gaps 11;

OY 5 NKLFFDLDAEENFVND-KFLKNELNDRRAQSQKD---KERDSQVIITDRLRLEERN 60
Db 852 DKLESLKDTBESNVLDIQRQLKAERETLKAMYSDALEPAQRELEINEDESELDERR 911
OY 61 ATTVSIIQ-----QALKAEMLCSTLKKOKMYLEOODDERKOAEAGRLRS 106
Db 912 LALENLDQMKSVKEKVADLEBELOEBOKLRNTLEKLKKRIEELFEENKRVDGSDSTIS 971
OY 107 KMKTEQIELLSQLPEVEEMI----RDGV-----GGSABOLAVCVS----- 148
Db 972 R--LEKIKDELQKVEELTBSFSDESKDKGLKTRVRLLSGELDDLVRIDSETRKSE 1028
OY 149 -----LKKEYINKERKA-SGEVADKLKRDLFSSSKSIQTYSILDQ-----AKLEIK 196
Db 1029 LIRQKKKEELEELKYQEALAAETAATKLAQE-AANKKLGQETTELNERFNSEVTARSHVE 1086
OY 197 SAQKDLOS----ADKEIMSIKKKLMQLQ 220
Db 1087 KSKKTLESOLVAVNNELBEEKKRRDALE 1114

RESULT 4
TANA_XENLA STANDARD: PRT: 1744 AA.
AC Q01550;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tanaibia.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tadpole head;
RX MEDLINE=92398961; PubMed=1524825;
RA Hemmati-Brivanlou A., Mann R.W., Harland R.M.;
RT "A protein expressed in the growth cones of embryonic vertebrate neurons defines a new class of intermediate filament protein.";
RL Neuron 9:417-428(1992).
CC -!- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.
CC -!- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURONA AND PERSIST DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL CORD.
CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outpost.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isdb-sib.ch/announce/
CC or send an email to license@sib.ch).
```

```

CC -----
DR EMBL: M99387; AAA49966.1; -
DR PIR: JH0720; JH0720.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 2.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone.
FT DOMAIN 1 12 HEAD.
FT DOMAIN 13 314 ROD.
FT DOMAIN 315 1744 TAIL.
FT DOMAIN 8 48 COIL 1A.
FT DOMAIN 49 60 LINKER 1.
FT DOMAIN 61 156 COIL 1B.
FT DOMAIN 157 179 LINKER 12.
FT DOMAIN 180 193 COIL 2A.
FT DOMAIN 194 199 LINKER 2.
FT DOMAIN 200 314 COIL 2B.
SQ SEQUENCE 1744 AA; 199561 MW; 6502EAC9FE6C4E93 CRC64;

```

```

Query Match 15.3%; Score 163.5; DB 1; Length 1744;
Best Local Similarity 25.8%; Pred. No. 0.058;
Matches 60; Conservative 61; Mismatches 73; Indels 39; Gaps 9;

```

```

QY 12 AQEENVLDREFLNELDNVRAQLSQDKERDSQVITDRLDTEERNATVSLQALG 71
   |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 33 ALEREN-----ELRKREIHSLSKSEKRCMKKHHEEM-KLRDALDDGHEMV----- 80
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 72 KAENLCSTLKKOMKYLEQODETKQAOEAGRLSKMKMTQETILLQSL-----P 123
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 81 QAEWVRDYSIEIEFYKORCLEEQKQAREDAKELSESKLLEETRAQIWKERLGQLEA 140
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 124 EEEEMIRMGQSAVEQLAVYCVSLKKEYENLEKARKA--SGEYADLRK--DLFSS-- 177
   |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 141 ELEILNRHEEKALMEB---EIASFSQRLNFRAVAPAFKPEVDYDARKLSEIWCQAV 197
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 178 ---RSKLTQVYSELDAQLELKSQKD-----LQSAKDEIMSLKKRLTMLQ 220
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 198 EEEKSEVSLBAGLSSEKSNLRYLEENKONRLQLQSLDKELVSLKMRKALE 250
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

RESULT 5
G160_MOUSE STANDARD; PRT; 1325 AA.
ID G160_MOUSE
AC P55937;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Golgin-160 (Male-enhanced antigen-2) (MEA-2).
GN GOLGA3 OR MEA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RX MEDLINE=97217683; PubMed=9063644;
RA Kondo M., Sutoh S.;
RT "Cloning and molecular characterization of cDNA encoding a mouse
RT male-enhanced antigen-2 (Mea-2): a putative family of the Golgi
RT autoantigen."
RL DNA Seq. 7:71-82(1997).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
CC TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY
CC DETECTABLE MALE ANTIGEN (SDM).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
CC FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN
CC LEYDIG CELLS, SPERMATOGENIA, OR SPERMATOCYTES.
CC -1- SIMILARITY: HIGH, TO HUMAN GOLGIN-160.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MEA-1, MEA-19 OR MEA-30 IS THE
CC INITIATOR.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: D78270; BAA19612.1; -
DR HSSP: P18852; ISCG.
DR MGD; MGI:96958; Golga3.
KW Spermatogenesis; Developmental protein.
FT DOMAIN 201 204 POLY-ALA.
SQ SEQUENCE 1325 AA; 149880 MW; 3230636962C687B0 CRC64;

```

```

Query Match 15.1%; Score 161; DB 1; Length 1325;
Best Local Similarity 23.1%; Pred. No. 0.06;
Matches 57; Conservative 48; Mismatches 90; Indels 52; Gaps 7;

```

```

QY 13 QEEENVLDREFLNELDNVRAQLSQDKERDSQVITDRLDTEERNATVSL 66
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 966 REHNSITETALAKREADVQLNLOVAVLQKREEDKQKOLVQALQSLKEMEVNSL 1025
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 67 QOALGKA-----EMLCSTLKKOMKYLEQO-----ODET 94
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1026 KEQMAARIEAGNRRHFKAATLELSEYKKELQAKENHVLQAEVDELQIDQDKHSGEI 1085
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 95 KOAOEAGRLSKMKMTQETILLQSLPEVEEMIRMGQSAVEQLAVYCVSLKKEYE 154
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1086 AQFQTELAERTQLO-----LQKRLK---DEQMSQPTGSQMEDLKMELDKREBIQ 1135
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 155 NLKEPARASGVADKLRKDLFFSSRSKLTQVYSELDAQLELKSQKDLQSAKDEIMSLK 214
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1136 SLKQDLDTQEG---KKELEGTQDTLTQITNSELMEVQEDLSFTQKXKFMQAKVSELKN 1192
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 215 KL-TMLQ 220
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 1193 NMKTLQ 1199
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

RESULT 6
MYS_AEOIR STANDARD; PRT; 1938 AA.
ID MYS_AEOIR
AC P24733;
DT 01-MAR-1992 (Rel. 21, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Myosin heavy chain, striated muscle.
OS Aequipecten irradians (Bay scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
OC Pectinoidea; Pectinidae; Argopecten.
OX NCBI_TaxID=31199;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=92011595; PubMed=1917970;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Complete primary structure of a scallop striated muscle myosin heavy
RT chain. Sequence comparison with other heavy chains reveals regions
RT that might be critical for regulation."
RL J. Biol. Chem. 266:18469-18476(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=91088319; PubMed=2263488;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
RT muscle myosin heavy chain."
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX MEDLINE=94173332; PubMed=8127365;

```

RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,  
 RA Szent-Gyorgyi A.G., Cohen C.;  
 RT "Structure of the regulatory domain of scallop myosin at 2.8-A  
 RT resolution.";  
 RL Nature 368:306-312(1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.  
 RX MEDLINE=96419133; PubMed=8805510;  
 RA Houdusse A., Cohen C.;  
 RT "Structure of the regulatory domain of scallop myosin at 2-A  
 RT resolution: implications for domain.";  
 RL Structure 4:21-32(1996).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE  
 CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X55714; CAA39247.1; -;  
 DR PIR: S13557; S13557.  
 DR PIR: A40997; A40997.  
 DR PDB: 1SCM; 30-APR-94.  
 DR PDB: 1WDC; 11-JUL-96.  
 DR InterPro: IPR000048; IO.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR002017; Spectrin.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IO; 2.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IO; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IO; 1.  
 KM Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KM ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.  
 FT DOMAIN 1 777  
 FT DOMAIN 778 805  
 FT DOMAIN 836 1938  
 FT DOMAIN 836 1938  
 FT NP\_BIND 176 183  
 FT NP\_BIND 176 183  
 FT MOD\_RES 693 693  
 FT MOD\_RES 703 703  
 FT MOD\_RES 778 821  
 FT HELIX 822 823  
 FT TURN 825 833  
 FT TURN 834 835  
 SQ SEQUENCE 1938 AA; 222821 MW; ASCECA127DIA4896 CRC64;  
 Query Match 15.1%; Score 160.5; DB 1; Length 1938;  
 Best Local Similarity 28.1%; Pred. No. 0.094;  
 Matches 56; Conservative 46; Mismatches 70; Indels 27; Gaps 7;  
 OY 10 DLADBEENVLDREFLKNEL-DNVR---AQLSOKREKRDSDVVIDTLRLTEERNATVVS 65  
 DB 1055 DLKSTQENVDELVEYKRELEENVRKREAEISLSNLSKLEDEONIVSOLQRIKEIQARIEE 1114

OY 66 LQALGKAEMLSTLKRMKYLEFOQODETKQAEAGRLRSKMTMEQIELLSQLEPV 125  
 DB 1115 LEELEFAERNRAKVEKORALRELEBELGERLDPEAGS-----ATSAQIELNKKRE-AEL 1168  
 OY 126 EEMTRDMGVGQSAVEQLAVCVSLKREYENLEKARKASGEADVKLRKDLFSSRSKLTQTV 185  
 DB 1169 LKIRRDLE---EASISQHQASALRKRHQD-----AANEMADGV-----DQLQKVK 1211  
 OY 186 SELDQAKLELSAKQDLOS 204  
 DB 1212 SKLEDKKRLKREMDLES 1230  
 RESULT 7  
 USOL\_YEAST STANDARD; PRT: 1790 AA.  
 AC P25386;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Intracellular protein transport protein USOL.  
 GN USOL OR INT1 OR YDL058W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180-1A;  
 RX MEDLINE=91185402; PubMed=2010462;  
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
 RA Yamasaki M.;  
 RT "A cytoskeleton-related gene, usol, is required for intracellular  
 RT protein transport in Saccharomyces cerevisiae.";  
 RL J. Cell Biol. 113:245-260(1991).  
 RN [2]  
 RP SEQUENCE OF 782-1790 FROM N.A.  
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,  
 RA Kendrick K.E.;  
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-8 FROM N.A.  
 RA Bai Y., Symington L.S.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
 CC COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
 CC ER AND THE GOLGI COMPLEX.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED  
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
 CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X54378; CAA38253.1; -;  
 DR EMBL: L03188; AAB00143.1; -;  
 DR EMBL: U53668; AAB66659.1; -;  
 DR PIR: A38455; A38455.  
 DR HSSP: P80220; IDIP.  
 DR SGD: S0002216; USOL.  
 DR InterPro: IPR002017; Spectrin.  
 KM Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.  
 FT DOMAIN 1 724  
 FT DOMAIN 725 1790  
 FT DOMAIN 465 487  
 CHARGED (HYPER-HYDROPHILIC).

FT	DOMAIN	991	1790	DISPENSABLE FOR THE PROTEIN FUNCTION.
FT	CONFLICT	1172	1786	ASP/GLU-RICH (ACIDIC).
FT	CONFLICT	847		G -> E (IN REF. 2).
FT	CONFLICT	924		E -> K (IN REF. 2).
FT	CONFLICT	1253		V -> I (IN REF. 2).
FT	CONFLICT	1319		I -> V (IN REF. 2).
FT	CONFLICT	1461		N -> S (IN REF. 2).
FT	CONFLICT	1581		G -> S (IN REF. 2).
FT	CONFLICT	1600		I -> V (IN REF. 2).
FT	CONFLICT	1661		R -> S (IN REF. 2).
FT	CONFLICT	1772		D -> DEEDDER (IN REF. 2).
SO	SEQUENCE	1790 AA;	206424 MW;	6CE2B216E9FD4818 CRC64;
Query Match 14.9%; Score 159; DB 1; Length 1790;				
Best Local Similarity 23.7%; Pred. No. 0.1;				
Matches 61; Conservative 57; Mismatches 85; Indels 54; Gaps				
QY	1	RTIINKLFFDLQ-----EENNVLDRELK-----NELDNVRAOLSKDKDKRDSYI	48	
DB	1369	RLLNLEGSSTIOEYSEKINTLEDELIRLQNONELKAKKIDWTRSEL-EKVSLSND----	1423	
QY	49	IDTLTDLERNATVVSLOAL-----GKAEMLCTPKKOMKYLEQOODETKOAGE	100	
DB	1424	-----ELLEKQNTIKSLQDELISKDKTTRNDKILSTERONKRPLESIKQLRAQGS	1478	
QY	101	AGRLSRKMTMQEIIQLLOSLPEVEEMIRMGVGSQAVE---QLAVYCSLKREYML	156	
DB	1479	KAKVEEGAKLKEESSEKAELEKSEMKKL---ESTIESNETELSSMETIRKSDKL	1535	
QY	157	KKARKASGVADAKLR-----KDLFSSSKQIQTYY---SELDQAKLELSKQ	199	
DB	1536	EQSKRSAEEDIKNIDHEKSDILSRINESEKDIIEELKSKLRIPAKSGSELETYKQELINNAQ	1595	
QY	200	KDLOSADKEIMSLKKKL	216	
DB	1596	EKIRINAEENTYLSKSL	1612	
RESULT 8				
MYHD_HUMAN	ID	MYHD_HUMAN	STANDARD:	PRT: 1938 AA.
AC	Q9UKX3:	Q95252:		
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Myosin heavy chain, skeletal muscle, extraocular (MyHC-eo).			
GN	MYH3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Extraocular muscle;			
RX	MEDLINE=99318869;	PubMed=10388558;		
RA	Weiss A., Schaffino S., Weinand L.A.;			
RT	"Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity.";			
RL	J. Mol. Biol.	290:61-75(1999).		
RN	[2]			
RP	SEQUENCE OF 1917-1938 FROM N.A.			
RC	TISSUE=Extraocular muscle;			
RX	MEDLINE=99026150;	PubMed=9806854;		
RA	Winters L.M., Briggs M.M., Schachaf F.;			
RT	"The human extraocular muscle myosin heavy chain gene (MYH13) maps to the cluster of fast and developmental myosin genes on chromosome 17.";			
RL	Genomics	54:188-189(1998).		
CC	-1- FUNCTION: MUSCLE CONTRACTION.			
CC	-1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC).			
CC	AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).			
CC	-1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.			

```

CC -I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -I- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -I- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEMOROUSIN (LMN) AND 1 HEAVY MEMOROUSIN (HMN). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AF111782; AAD29948.1; -
CC EMBL, AF075248; AAC83241.1; -
CC HSSP, P08799; 1MND.
CC MIM, 603487; -
CC InterPro: IPR000048; IQ.
CC InterPro: IPR004009; myosin_N.
CC InterPro: IPR002928; myosin_tail.
CC InterPro: IPR001609; myosin_head.
CC Pfam: PF00612; IQ; 2.
CC Pfam: PF00063; myosin_head; 1.
CC Pfam: PF02736; myosin_N; 1.
CC Pfam: PF01576; myosin_tail; 1.
CC PRINTS: PR00193; MYOSINHEAVY.
CC Prodom: PD000355; myosin_head; 1.
CC SMART: SM00015; IQ; 1.
CC SMART: SM00242; MYSC; 1.
CC PROSITE: PS50096; IQ; 1.
CC KW myosin; muscle protein; coiled coil; thick filament; Actin-binding;
CC Calmodulin-binding; ATP-binding; Methylation; Alkylation;
CC Multigene family.
CC FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
CC FT NP-BIND 843 1938 COILED COIL (POTENTIAL).
CC FT DOMAIN 179 186 ATP (POTENTIAL).
CC FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).
CC FT DOMAIN 751 775 ACTIN-BINDING (BY SIMILARITY).
CC FT MOD_RES 130 130 METHYLATION (PT1-) (POTENTIAL).
CC FT MOD_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).
CC FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
CC SQ SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CD5 CRC64;
Query Match 14.9%; Score 159; DB 1; Length 1938;
Best Local Similarity 23.3%; Pred. No. 0.11;
Matches 57; Conservative 48; Mismatches 86; Indels 54; Gaps 8;
QY 13 GREENVL-----DREFLNKELNDNVRAQLSQKDKER-----RDSQVLIIDTLRDTLLEERN 60
DB 1296 EKESKLSISLQFSKQALTPQOLELKRQMEERFKANNAHMAHLOSSRHODDLREQYEERQ 1355
QY 61 ATTVSLQOALGKAE-----MLCSTLKQKKYLEQOODETKQAQAEAGRLR 105
DB 1356 EAKAELEQRLAKSKANSEVAQMKTKYETDAIQTFEELAEAKKKLAQRLOEAEKTEETA--N 1412
QY 106 SAKMKMEQELLLOLQDLPVEEMINDMGVGSANVQQLAVYCSLAKKEYENLK----- 157
DB 1413 SKCALETKQRLOG---EVEDLMDL-----FRSHTACTPLDKKQRFNFDKYLAEWKQ 1462
QY 158 --EARKAGEVNDK----LRKDLFFSSRSKLTQVYSELDOAKLELKSQAOKDLSADKEIMS 211
DB 1463 KLDSEQAELEAAQKSRSLSTELFKMRNNAEEVYQOLELRREKNKNDLEISDLTEQIAE 1522
QY 212 LKRL 216

```

Db 1523 TGNL 1527

RESULT 9

MYH7\_RAT STANDARD: PRT; 1935 AA.

AC P02564;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, cardiac muscle beta isoform (MyHC-beta).

GN MYH7.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RA MEDLINE=90016823; PubMed=2798112;

RT Kraft R., Bravo-Zehnder M., Taylor D., Leinwand L.A.;

RT "Complete nucleotide sequence of full length cDNA for rat beta

RL cardiac myosin heavy chain.";

RL Nucleic Acids Res. 17:7529-7530(1989).

RN (2)

RP DISCUSSION OF SEQUENCE.

RX MEDLINE=90133919; PubMed=2614840;

RA McNally E.M., Kraft R., Bravo-Zehnder M., Taylor D., Leinwand L.A.;

RT "Full-length rat alpha and beta cardiac myosin heavy chain sequences.

RT Comparisons suggest a molecular basis for functional differences.";

RL J. Mol. Biol. 210:665-671(1989).

RN (3)

RP SEQUENCE OF 1524-1935 FROM N.A.

RX MEDLINE=62220036; PubMed=7045682;

RA Mahdavi V., Periasamy M., Nadal-Ginard B.;

RT "Molecular characterization of two myosin heavy chain genes expressed

RT in the adult heart.";

RL Nature 297:659-664(1982).

RN (4)

RP SEQUENCE OF 1871-1935 FROM N.A.

RC STRAIN=Wistar; TISSUE=Heart;

RX MEDLINE=65179510; PubMed=6241892;

RA Mahdavi V., Lompre A.M., Chambers A.P., Nadal-Ginard B.;

RT "Cardiac myosin heavy chain isozymic transitions during development

RT and under pathological conditions are regulated at the level of mRNA

RT availability.";

RL Eur. Heart J. 5:181-191(1984).

CC -1- FUNCTION: MUSCLE CONTRACTION.

CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2

CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)

CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING

CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY

CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT

CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE

CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED

CC SUBFRAGMENT (S2).

CC -1- MISCELLANEOUS: THE CARDIAC ALPHA ISOFORM IS A 'FAST' ATPASE

CC MYOSIN, WHILE THE BETA ISOFORM IS A 'SLOW' ATPASE.

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC -----

CC EMBL: X15939; CAA34065.1; -;

CC DR EMBL: J00752; AAA41654.1; -;

CC DR EMBL: M32698; AAA41659.1; -;

CC DR PIR: S06006; S06006.

CC DR PIR: A02989; A02989.

CC DR HSSP: P08799; 1MMD.

CC DR InterPro: IPR000048; IQ.

CC DR InterPro: IPR004003; Myosin\_N.

CC DR InterPro: IPR002928; Myosin\_Tail.

CC DR InterPro: IPR001609; myosin\_head.

CC DR Pfam: PF00612; IQ; 2.

CC DR Pfam: PF00063; Myosin\_head; 1.

CC DR Pfam: PF02736; Myosin\_N; 1.

CC DR Pfam: PF01576; Myosin\_Tail; 1.

CC DR PRINTS: PR00193; MYOSINHEAVY.

CC DR PRODOM: PD000355; myosin\_head; 1.

CC DR SMART: SM00015; IQ; 1.

CC DR SMART: SM00242; MYSC; 1.

CC DR PROSITE: PS50096; IQ; 1.

CC KW Myosin; muscle protein; coiled coil; thick filament; Actin-binding;

CC KW ATP-binding; Methylation; Alkylation; Multigene family;

CC KW Calmodulin-binding.

CC MW 1 780

CC FT DOMAIN 781 810

CC FT NP\_BIND 840 1935

CC FT NP\_BIND 178 185

CC FT DOMAIN 655 677

CC FT DOMAIN 757 771

CC FT MOD\_RES 129 129

CC FT MOD\_RES 695 695

CC FT MOD\_RES 705 705

CC FT MOD\_RES 1529 1531

CC FT CONFLICT 1731 1731

CC FT CONFLICT 1784 1784

CC FT CONFLICT 1851 1851

CC FT CONFLICT 1858 1858

CC FT SEQUENCE 1935 AA; 223082 MW; C8376C324A7BD82B CRC64;

Query Match 14.9%; Score 158.5; DB 1; Length 1935;

Best Local Similarity 20.8%; Pred. No. 0.12;

Matches 62; Conservative 60; Mismatches 87; Indels 89; Gaps 9;

QY 1 RTITNKLFDDAQBENVLDR-----EFLK----- 25

Db 1129 RAKVEKLRSDLSRELESEIRLEBAGATSVQIEMKKREAEPQMRDRLEATLQHEAT 1188

QY 26 -----NELDVNRQLSQDKKEDRSQVITDLDPTLEERNATVVSLOQA 69

Db 1189 AALRRKKAADVSLGEIDVLRQKQLEKESFKLELDVSNMQ-----IIRA 1241

QY 70 LGKAEMLCSTLKQMKYLEQOODETKQAQEFAGRLSKMKT-----MQEILLQSQ 121

Db 1242 KANLEKMQRTLEDQNMNRSKAEEQSVNDLTRRAQLQFENGLSNQDKEKALI-SQ 1300

QY 122 LP-----EVEEMIRDMGVGASAVEQLAVCYSLKKEYENLKEARKASCEV 166

Db 1301 LTRGRLTYTQOLEDLKRLQLEEVKAKNALAHALASARHCDCLRLQRYEEFEAAAELORV 1360

QY 167 ADKLRLKDLSSRSKLOT---VYSFLDQAKLELSAQKDLDSAPKEIMSLKKKLMLQ 220

Db 1361 LSKANSEVAQMRFTYETDAIQRTLELEBAKKKL--AQR-LQDAEAEVAEVAKKCSLE 1415

RESULT 10

MYSS\_CYPCA STANDARD: PRT; 1935 AA.

AC Q90339;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, fast skeletal muscle.  
 OS *Cyprinus carpio* (Common carp).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 CC Cypriniformes; Cyprinidae; Cyprinus.  
 CX NCBI\_Taxid=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fast muscle;  
 RX MEDLINE=97352533; PubMed=9208928;  
 RA Hirayama Y., Watabe S.;  
 RT "Structural differences in the crossbridge head of temperature-associated myosin subfragment-1 isoforms from carp fast skeletal muscle.";  
 RT Eur. J. Biochem. 246:380-387(1997).  
 RN [2]  
 RP SEQUENCE OF 981-1935 FROM N.A.  
 RC TISSUE=Fast muscle;  
 RX MEDLINE=97176447; PubMed=9023993;  
 RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;  
 RT "cDNA cloning of myosin heavy chain isoforms from carp fast skeletal muscle and their gene expression associated with temperature acclimation.";  
 RT J. Exp. Biol. 200:27-34(1997).  
 RN [3]  
 RP SEQUENCE OF 1387-1528 FROM N.A.  
 RX MEDLINE=95194396; PubMed=7887920;  
 RA Watabe S., Imai J., Nakaya M., Hirayama Y., Okamoto Y., Masaki H., Uozumi T., Hirono I., Aoki T.;  
 RT "Temperature acclimation induces light meromyosin isoforms with different primary structures in carp fast skeletal muscle.";  
 RT Biochem. Biophys. Res. Commun. 208:118-125(1995).  
 CC -1 FUNCTION: MUSCLE CONTRACTION.  
 CC -1 SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1 SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1 DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1 PRIM: TWO CYSTEINE RESIDUES IN THE SI DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1 MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).  
 CC -1 SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@sib-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch)).  
 CC -----  
 DR EMBL: D89992; BAA22069.1; -;  
 DR EMBL: D50476; BAA09069.1; -;  
 DR EMBL: D43700; BAA07802.1; -;  
 DR HSSP: P08799; IAMD.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_N.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000335; myosin\_head; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.

DR PROSITE: PS50096; IQ; 1.  
 KW ATP-binding: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding: Methylation; Alkylation; Calmodulin-binding;  
 KW Multigene family.  
 KW DOMAIN 1 MYOSIN HEAD-LIKE.  
 FT DOMAIN 782 811 IQ.  
 FT DOMAIN 812 839 HINGE.  
 FT DOMAIN 840 1935 COILED COIL (POTENTIAL).  
 FT NP\_BIND 178 185 ATP (POTENTIAL).  
 FT DOMAIN 659 681 ACTIN-BINDING.  
 FT DOMAIN 761 775 ACTIN-BINDING.  
 FT MOD\_RES 129 129 METHYLATION (TR-1) (POTENTIAL).  
 FT MOD\_RES 699 699 ALKYLATION (SH-1).  
 FT MOD\_RES 709 709 ALKYLATION (SH-2).  
 SQ SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;  
 Query Match 14.7%; Score 157; DB 1; Length 1935;  
 Best Local Similarity 22.4%; Pred. No. 0.14;  
 Matches 57; Conservative 51; Mismatches 87; Indels 60; Gaps 8;  
 QY 15 EENVLDREF-----LKNELDVNRQLSKDKREKRSQYIIDLRPTLEER 59  
 DB 1180 EESTLQHEATTAALRKQADSVAELGEQIDMLQRYKQLEKSEYKMEIDLTSMNE-- 1237  
 QY 60 NATVSLQALGKAEMLCSFLKKQKYLEQOODETKQAEGRLSKMT-----M 111  
 DB 1238 -----AVAKAKANLEKMTLEDLSLTKSDENVRLNDMNRALQTENGEFSROL 1292  
 QY 112 EQIELLIQSOLP-----EVEEMIRMGVGSAGVEQLAVYCVSLKREYNL 156  
 DB 1293 EEKEALV-SQLTRGQAVTQOIEELKRHIEEVRKAKNALAHVASARHCDLREQYEE 1351  
 QY 157 KEAR-----KASGEVAD---KIKRDLFSSRSKQLQVVEILDQAKELKSAQKDQSA 205  
 DB 1352 QEAKELDRGMSKANSEVAQMRTYETAIOFTELEBAKKRLAQ---RLDDEESIEAV 1408  
 QY 206 DKEIMSLKKLTMLQ 220  
 DB 1409 NSKASLEKTRQRLQ 1423  
 RESULT 11  
 VDP\_MOUSE STANDARD; PRT; 941 AA.  
 AC 09:120:  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE General vesicular transport factor p115 (Transcytosis associated protein) (TAP) (Vesicle docking protein) (Fragment).  
 GN VDP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Thymus;  
 RA Han S.-J.;  
 RT "Identification of mouse TAP (transcytosis associated protein/p115)."  
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1 FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR INTERCISTERNAL TRANSPORT IN THE GOLGI STACK. IT IS REQUIRED FOR TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR AND TARGET MEMBRANS IN PROXIMITY (BY SIMILARITY).  
 CC -1 SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE (BY SIMILARITY).  
 CC -1 DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.



CC -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER;  
CC PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.  
CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLDI MEMBRANE;  
CC PHOSPHORYLATION PROMOTES DISSOCIATION (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF096868; AAC72967.1; -  
DR MGI: 1929095; VAP.  
DR InterPro: IPR000225; Armadillo.  
DR InterPro: IPR000219; Rhogef.  
DR SMART: SM00325; Rhogef. 1.  
DR PROSITE: PS50176; ARM\_REPEAT; UNKNOWN\_1.  
KW Transport: Protein transport; Golgi stack; Membrane; Coiled coil;  
KW Phosphorylation.  
FT NON\_TER 1  
FT DOMAIN <1 619 GLOBULAR HEAD.  
FT DOMAIN 620 912 COILED COIL (POTENTIAL).  
FT DOMAIN 917 941 ASP/GLU-RICH (ACIDIC).  
FT MOD\_RES 922 922 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 941 AA; 105152 MW; AB5C10855CD7E508 CRC64;

Query Match 14.6%; Score 155.5; DB 1; Length 941;  
Best Local Similarity 19.5%; Pred. No. 0.084;  
Matches 64; Conservative 52; Mismatches 81; Indels 131; Gaps 8;

QY 15 EENVLDREFLK--NELDNV--RAQLSQDKERKDSQVYIIDLRLDLEERNATVYSLQAL 70  
DB 587 EYMFIDHFTFLVKLEGVITKALYKSSSEKKEEV-----KKTLEOHNDINVTYKMMI 641  
QY 71 GKAEMLCSTLKKOMKYLBOODE-----TKQAQ----- 98  
DB 642 REQDLQLEELKQVSTLKQCNQLOLAVTQAASQIQKHQDYQNLKVLQKDNHQQSHG 701  
QY 99 -----EAGRLRSKMTMEQIELLLSQLPVEEMIRDMGVGSA--VEQLAVYC 146  
DB 702 DGAQVNGIOPFELISRLREIELEKSQLLOGQLAEKDSLLENLSQASGSEQASATC 761  
QY 147 ----- 146  
DB 762 PPDPDQVAVELKQELTALKSQLCSQSLBITRLQTEKNCILLQRAETLAKSVPEGESEHVS 821  
QY 147 -----VSLKKEYENLKEARKASGEVADKLKDLFSRSKLQTVYSLDQAKTEL 195  
DB 822 AAKTTDVGRUSALLQETKELKNEIKALSEERTATQKQDSSNSTAITALQTEKDLDEV 881  
QY 196 KSAQKD-----LQSADKEIMSLKKL 216  
DB 882 TDSKKEQDLDLVLLADPDQKILSLKSL 909

RESULT 12  
XCP\_E\_XENIA  
ID XCP\_E\_XENIA STANDARD: PRT; 1203 AA.  
AC P50533:  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Chromosome assembly protein XCAP-E.  
GN XCAP-E.  
OS Xenopus laevis (African clawed frog).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
CC Xenopodinae; Xenopus.  
NCBI\_TaxID=8355;

RN [1]  
CC SEQUENCE FROM N.A.  
RX MEDLINE-95042742; PubMed-7954811;  
RA Hirano T., Mitchison T.J.;  
RT "A heterodimeric coiled-coil protein required for mitotic chromosome  
RL condensation in vitro.";  
RL Cell 79:449-458(1994).  
CC -1- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF  
CC MITOTIC CHROMOSOMES.  
CC -1- SUBUNIT: ASSOCIATES WITH XCAP-C PROBABLY AS HETERODIMER.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR. DURING CHROMOSOME ASSEMBLY IN  
CC MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND  
CC FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.  
CC -1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS.  
CC -1- PLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.  
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U13674; AAA64680.1; -  
DR HSSP: P07751; 1A73  
DR InterPro: IPR003405; SMC\_C.  
DR InterPro: IPR003395; SMC\_N.  
DR Pfam: PF02483; SMC\_C; 1.  
DR Pfam: PF02463; SMC\_N; 1.  
KW Mitosis; ATP-binding; Coiled coil; Nuclear protein.  
FT NP\_BIND 32 39  
FT DOMAIN 172 356 ATP (POTENTIAL).  
FT DOMAIN 397 513 COILED COIL (POTENTIAL).  
FT DOMAIN 670 1032 COILED COIL (POTENTIAL).  
FT DOMAIN 1095 1121 ALA/ASP-RICH (DA-BOX).  
SQ SEQUENCE 1203 AA; 136341 MW; 04323DD027DF309 CRC64;

Query Match 14.6%; Score 155.5; DB 1; Length 1203;  
Best Local Similarity 22.0%; Pred. No. 0.11;  
Matches 48; Conservative 54; Mismatches 87; Indels 29; Gaps 5;

QY 16 ENVIDR-----EFLNELDNVRAQLSQDKERKDSQVYIIDLRLDLEERNATVYSLQ 67  
DB 707 KNTYERYQLKQOMKEMKSEAEELLQTKLQSSYHKOQEE--LDSLKQYTESSEETLKNTK 764  
QY 68 QALGKAEMLCSTLKKOMKYLBOQOD-ETKQAQEEAGRLRSKM---KTMQIELLLSQSL 122  
DB 765 EYQKKAEEKFKVLEHKMNAEARELEKQAKLDTKAKKADASNKKKKEQGEVDALV 824  
QY 123 PEVEEMIRDMGVGQSAVOLAVYCVSLKKEYENLKEARKASGEVADKLKDLFSRSKLQ 182  
DB 825 LELFELKREQ-----TTYKQIETVDEAMKAVQEQDMSADEVSKKKEAVK 870  
QY 183 TVYSELQAKLEKSAQKDLSADKEIMSLKKKLTMLQ 220  
DB 871 KAQDELAKQKELIMGHDKREIKTSSEAGKLRNNNDLQ 908

RESULT 13  
MYH4\_RABIT  
ID MYH4\_RABIT STANDARD: PRT; 1938 AA.  
AC Q28641:  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin heavy chain, skeletal muscle, juvenile.  
OS Oryctolagus cuniculus (Rabbit).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
NCBI\_TaxID=9986;

```

RN      [1]
SEQUENCE FROM N.A.
RC      STRAIN-NEW ZEALAND WHITE: TISSUE-Skeletal muscle;
RA      Maeda K., Hoshinova E., Roesch-Kleinkauf A., Schuster H., Gasperik J.,
RT      Wittinghofer A.;
RT      "Isolation, sequencing of myosin heavy chain cDNA from rabbit
RT      skeletal muscle and a novel cosynthesis of S-1 fragment with the
RT      essential and regulatory light chains.";
RL      Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
CC      -1- FUNCTION: MUSCLE CONTRACTION.
CC      -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC      HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC      AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC      -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC      -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC      CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC      CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC      -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC      MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC      SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC      SUBFRAGMENT (S2).
CC      -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U32574; AAA74199.1; -.
DR      HSSP: P08799; 1MMD.
DR      InterPro: IPR000048; IQ.
DR      InterPro: IPR004009; Myosin_N.
DR      InterPro: IPR002928; Myosin_tail.
DR      InterPro: IPR001609; myosin_head.
DR      Pfam: PF00612; IQ: 2.
DR      Pfam: PF00063; myosin_head; 1.
DR      Pfam: PF02736; myosin_N; 1.
DR      Pfam: PF01576; Myosin_tail; 1.
DR      PRINTS: PR00193; MYOSINHEAVY.
DR      PRODOM: PD000355; myosin_head; 1.
DR      SMART: SM00015; IQ: 1.
DR      SMART: SM00242; MYSC: 1.
DR      PROSITE: PS50096; IQ: 1.
KW      Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW      Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW      Multigene family.
FT      DOMAIN 1 783 813 MYOSIN HEAD-LIKE.
FT      DOMAIN 2 842 842 10.
FT      NP_BIND 179 186 COILED COIL (POTENTIAL).
FT      DOMAIN 3 658 680 ACTIN-BINDING (BY SIMILARITY).
FT      DOMAIN 4 760 774 ACTIN-BINDING (BY SIMILARITY).
FT      MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
FT      MOD_RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).
FT      MOD_RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).
FT      MOD_RES 756 756 METHYLATION (SH-1) (BY SIMILARITY).
FT      MOD_RES 698 698 ALKYLATION (SH-2) (BY SIMILARITY).
FT      MOD_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
SQ      SEQUENCE 1938 AA; 223064 MW; DBA8A2EC5B182626 CRC64;

```

Query Match 14.5%; Score 155; DB 1; Length 1938;  
 Best Local Similarity 23.8%; Pred. No. 0.19;  
 Matches 54; Conservative 49; Mismatches 80; Indels 44; Gaps 7;

Oy 24 LKNEIDNRAQLSKDKKRDQVLIIDRLRLERNAIVSLQALKAEMLCSTLKKQ 83  
 Db 1206 LGEQIDNLRVKKLEKSEKSEIMEDDLASNM-----TVSKAKGLMEKMCRTLEDQ 1258

```

Oy      84 MKYLEQODETK-----QAQEEAGRL-----RSKMTMEQIELL 117
Db      1259 VSELKTKKEENHRLINDLSAQARQLQTESRQDLDEKDSLVLSKQKAFQIQIEL 1318
Oy      118 LOSQLEPEVEMIRDMGVQSAVEQLAVYCVSIRKEYENLKEARKASGEVADKLKRDLFSS 177
Db      1319 KR-----QLEEFIKAKSALAHALQSAHRHCDLLREQYEEQEKALQIRAMSKANSEVQW 1374
Oy      178 RSKLQTF-----VSELDQAKLEKSNOKDQSDREKMSLKKKLTWMLQ 220
Db      1375 RTKYETDAIQRTEELEAKKTL--AQR-LDQAEHVEAVNAKCASTLE 1418

RESULT 14
MYH3_RAT
ID      MYH3_RAT      STANDARD:      PRT: 1940 AA.
AC      P12847.
DT      01-OCT-1989 (Rel. 12, Created)
DR      01-OCT-1989 (Rel. 12, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Myosin heavy chain, fast skeletal muscle, embryonic.
GN      MYH3.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX      NCBI_Taxid=10116.
RN      [1]
RP      SEQUENCE FROM N.A. Pubmed=3783701;
RX      MEDLINE=87060988;
RA      Streicher E.E., Streicher-Page M.-A., Perriard J.C., Perlasamy M.,
RA      Nadal-Ginard B.;
RT      "Complete nucleotide and encoded amino acid sequence of a mammalian
RT      myosin heavy chain gene. Evidence against Intron-dependent evolution
RT      of the rod.";
RL      J. Mol. Biol. 190:291-317(1986).
CC      -1- FUNCTION: MUSCLE CONTRACTION.
CC      -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC      HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC      AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC      -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC      -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC      CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC      CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC      -1- PFM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC      ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC      -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC      MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC      SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC      SUBFRAGMENT (S2).
CC      -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X04267; CAA27817.1; -.
DR      PIR: A24922; A24922.
DR      HSSP: P08799; 1MMD.
DR      InterPro: IPR000048; IQ.
DR      InterPro: IPR004009; Myosin_N.
DR      InterPro: IPR002928; Myosin_tail.
DR      InterPro: IPR002017; Spectrin.
DR      InterPro: IPR001609; myosin_head.
DR      Pfam: PF00612; IQ: 2.
DR      Pfam: PF00063; myosin_head; 1.
DR      Pfam: PF02736; myosin_N; 1.
DR      Pfam: PF01576; Myosin_tail; 1.
DR      PRINTS: PR00193; MYOSINHEAVY.

```

DR ProDom; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KW Multigene family.  
 FT DOMAIN 1 781 MYOSIN HEAD-LIKE.  
 FT DOMAIN 782 811 IQ.  
 FT DOMAIN 840 1933 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT DOMAIN 656 678 ACTIN-BINDING.  
 FT DOMAIN 758 772 ACTIN-BINDING.  
 FT MOD\_RES 130 130 METHYLATION (TR-1) (POTENTIAL).  
 FT MOD\_RES 696 696 ALKYLATION (SH-2).  
 FT MOD\_RES 706 706 ALKYLATION (SH-2).  
 SQ SEQUENCE 1940 AA; 223857 MW; B5D546A59653A696 CRC64;

Query Match 14.5%; Score 155; DB 1; Length 1940;  
 Best Local Similarity 23.8%; Pred. No. 0.19; Mismatches 83; Indels 36; Gaps 7;  
 Matches 53; Conservative 51;

QY 24 LKNELDNVRALQSOKREKRDSDVIIDLTLDTEERNATVVSLOALGKAEMLCSTLRKQ 83  
 DB 1204 LAEQIDMLQRYKQKLEKESEFKLEIDDLSSVSE-----SVSKSKANLEKICRTLEDQ 1256  
 QY 84 MKYLEQOODETKQA-----QEEAGRLRSKMTMEQIELL-----QSOLPEVDEM 128  
 DB 1257 LSEARGNEETQRLSLETTQKSRLOTEAGELSRQLEEKESIVSLSRSKQATFOQIEEL 1316  
 QY 129 IRDWGQSGAVESQAVLV-----CVSLKREYENLKEARKASGVADLRKDLFSSRSKL 181  
 DB 1317 KRLQEEENKAKNALAHALQSRHDCDLLRDYEEQSGKAELOALSKANSEVAVQMKTKY 1376  
 QY 182 QT-----VYSELDAQLELKSQKDLQSDAKREIMSLKKRLTMLQ 220  
 DB 1377 ETDALQRTLEELKAKRL--AQR-LQOSEQVEAVNAKKCASTLE 1416

RESULT 15  
 CENT\_HUMAN STANDARD; PRT; 3210 AA.  
 ID CENT\_HUMAN P49454; Q13246; Q13171;  
 AC P49454; Q13246; Q13171;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE CNP-F, kinetochore protein (Centromere protein F) (Mitosis) (AH antigen).  
 GN CNP-F.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast carcinoma;  
 RA MEDLINE=95348175; Pubmed=7542657;  
 RA Liao H., Winkler R.J., Mack G., Rattner J.B., Yen T.J.;  
 RT "CNP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis.";  
 RL J. Cell Biol. 130:507-518(1995).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95379848; Pubmed=7651420;  
 RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B.,  
 RA Jones D., Yang-Feng T.L., Lee W.-H.;  
 RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein that is specifically involved in mitotic-phase progression.";  
 RL Mol. Cell. Biol. 15:5017-5029(1995).  
 RP SEQUENCE OF 2194-3210 FROM N.A.  
 RX MEDLINE=95336446; Pubmed=7612011;

RA Li Q., Ke Y., Kapp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;  
 RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal domain sufficient for nuclear localization.";  
 RL Biochem. Biophys. Res. Commun. 212:220-228(1995).  
 RN [4]  
 RP CHARACTERIZATION.  
 RA MEDLINE=95370296; Pubmed=7642639;  
 RX Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;  
 RT "The C terminus of mitosis is essential for its nuclear localization, centromere/kinetochore targeting, and dimerization.";  
 RL J. Biol. Chem. 270:19545-19550(1995).  
 RN [5]  
 RP CHARACTERIZATION.  
 RA MEDLINE=98437347; Pubmed=9763420;  
 RX Chan G.K.T., Schaar B.T., Yen T.J.;  
 RT "Characterization of the kinetochore binding domain of CNP-E reveals interactions with the kinetochore proteins CNP-F and hBUBR1.";  
 RL J. Cell Biol. 143:49-63(1998).  
 CC -1- FUNCTION: PROBABLY REQUIRED FOR KINETOCHORE FUNCTION, INVOLVED IN CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH RETINOBLASTOMA PROTEIN (RB), CNP-E AND BUBR1.  
 CC -1- SUBUNIT: HOMO- OR HETERODIMER.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS), REORGANIZATION TO THE KINETOCHORE/CENTROMERE (CORONAL SURFACE OF THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.  
 CC -1- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.  
 CC -1- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U19769; AAA82889.1; -;  
 DR EMBL; U30872; AAA82935.1; -;  
 DR EMBL; U25725; AAA86889.1; -;  
 DR HSSP; P02649; 1LE4.  
 DR MIM; 600236; -;  
 KW Chromosomal protein; Nuclear protein; Centromere; Coiled coil;  
 KM Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat.  
 FT DOMAIN 14 197  
 FT DOMAIN 273 769  
 FT DOMAIN 823 1328  
 FT DOMAIN 1642 1746  
 FT DOMAIN 1862 2987  
 FT DOMAIN 2207 2568  
 FT REPEAT 2207 2386  
 FT REPEAT 2389 2568  
 FT DOMAIN 3015 3032  
 FT CONFLICT 16 16  
 FT CONFLICT 250 250  
 FT CONFLICT 272 272  
 FT CONFLICT 611 611  
 FT CONFLICT 1494 1589  
 FT CONFLICT 1611 1611  
 FT CONFLICT 1811 1811  
 FT CONFLICT 2242 2243  
 FT CONFLICT 2335 2335  
 FT CONFLICT 2492 2492  
 FT CONFLICT 2545 2561  
 SQ SEQUENCE 3210 AA; 367589 MW; 11D83524960E4334 CRC64;

Query Match 14.5%; Score 155; DB 1; Length 3210;  
 Best Local Similarity 23.6%; Pred. No. 0.31; Mismatches 90; Indels 40; Gaps 7;  
 Matches 57; Conservative 55;  
 QY 19 LDREKNELDNVRALQSOKREKRDSDVIIDLTLDTEERNATVVSLOALGKAEMLCSTLRKQ 77  
 DB 1204 LAEQIDMLQRYKQKLEKESEFKLEIDDLSSVSE-----SVSKSKANLEKICRTLEDQ 1256

```

Db 2299 LDVLTSEKENLTKQIOEKOGOLSELDKLSFKSLKEKEQAEIQIKESKTAVEMIQ 2358
Oy 78 STLKQO-----MKYLPQOOD---ETKQAGEAGRLRSKMKTKMOTELLIQS 120
Db 2359 NQKEELNVAVALGDDGEIMKATEOSLDPRPEEHQOLNLSIEKRLARLEADEKQOLCYIQ 2418
Oy 121 QLPPE-----VEEMIRDMGVGOSAVBQOLAVYCVSLKREYENLKEARKASGEVADK 169
Db 2419 QKRESEHHADLLKGRVENLEHERLEFIARTNOHALLEAVNSGVEYTLKAKTEGTMQSLRG 2478
Oy 170 LRKLFSSRSKLQTVYSELDQ-----AKLEL--KSQKDLQSDAKELMSLKK-----LTM 218
Db 2479 LELDVLTIRSEKEDLTNELQKOEERISELEIINSSFENILQEKDEQVKMEKESSTAMEM 2538
Oy 219 LQ 220
Db 2539 LQ 2540

RESULT 16
VDP_RAT STANDARD: PRT: 959 AA.
AC P1542:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE General vesicular transport factor p15 (Transcytosis associated
protein) (TAP) (Vesicle docking protein).
CN VDP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95132632; PubMed=7831323;
RA Barroso M., Nelson D.S., Szul E.;
RT "transcytosis-associated protein (TAP)/p15 is a general fusion
factor required for binding of vesicles to acceptor membranes.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:527-531(1995).
CC -1- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR
INTERCISTERNAL TRANSPORT IN THE GOLGI STACK; IT IS REQUIRED FOR
TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO
THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY
INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR
AND TARGET MEMBRANES IN PROXIMITY.
CC -1- SUBUNIT: HOMODIMER WITH TWO N-TERMINAL HEADS AND A C-TERMINAL
COILED-COIL TAIL. DIMER FORMED BY PARALLEL ASSOCIATION OF THE
TAILS.
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES
BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.
CC -1- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-
COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
CC -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER.
CC PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.
CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE;
CC PHOSPHORYLATED PROMOTES DISSOCIATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE VDP/OSOI/YELO4 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14192; AAC62632.1; -.
DR EMBL: U15589; AAC52151.1; -.
DR InterPro: IPR000225; Armadillo.
DR PROSITE: PS0176; ARM_REPEAT; UNKNOWN_1.
KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil;
KW Phosphorylation.
FT DOMAIN 1 637 GLOBULAR HEAD.
FT DOMAIN 638 930 COILED COIL (POTENTIAL).
FT DOMAIN 935 959 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 940 940 PHOSPHORYLATION (BY SIMILARITY).
FT CONFLICT 591 591 S -> P (IN REF. 2).
FT CONFLICT 658 658 M -> V (IN REF. 2).
FT CONFLICT 816 816 S -> R (IN REF. 2).
FT CONFLICT 873 873 A -> S (IN REF. 2).
SQ SEQUENCE 959 AA; 107162 MW; 356394B48C7E003B CRC64;

Query Match 14.5%; Score 154.5; DB 1; Length 959;
Best Local Similarity 20.1%; Pred. No. 0.097;
Matches 66; Conservative 53; Mismatches 78; Indels 131; Gaps 9;

Oy 15 EENVLDREFLK--NELDNV--RAQLSQDKKRDQOYIPTLRDTLEERNATVYSIQAL 70
Db 605 EYMLFDHEFTKLVKELEGVITKALTKSEDEKKEEV-----KTLQHDNIYVHYKNMI 659
Oy 71 GKAEMLCSTLKKMKKYLEQODE-----TKQAO----- 98
Db 660 REQLQLEELKQVSTLKQNEQVATQOASQIQGHKQDYNLKVQLGKDNHQSLS 719
Oy 99 -----PEAGRLSKSKMTMEQIELLQSQLPYVEEMTRDM-----GVGSAV--- 139
Db 720 DGAQVNGIQPEISRLREELRLSHQVLLQSQLAERDVIENLRSSQVSGMSBQALATC 779
Oy 140 -----EQLAVY----- 145
Db 780 SPRADQVAVELKQELSLKLSQCSQSLIEITRLQENSELDQRAATLAKSVPEGESELYT 839
Oy 146 -----CVSLKREYENLKEARKASGEVADKLKRDLFSSRSKLQTVYSELDQKLEL 195
Db 840 AAKTVDVEGRSLALQETKELKNEIKALSERRTAIQOLDSSNSTIAILOTEDKLYLEV 899
Oy 196 KSQKQD-----LQSADEKELMSLKKRL 216
Db 900 TDSKKEQDDLVLADQDKILSLKRL 927

RESULT 17
MYSN_ACACA STANDARD: PRT: 1509 AA.
AC P05659;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin II heavy chain, non muscle.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_Taxid=5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87308395; PubMed=3040773;
RA Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT "Complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail.";
RL J. Cell Biol. 105:913-925(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI

```

CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE THE REGULATORY PHOSPHORYLATION SITES RESIDE.

CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 IO DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: Y00624; CAA68663.1; -

DR PIR: A27224; A27224.

DR HSSP: P08799; 1MND.

DR Interpro: IPR000046; IQ.

DR Interpro: IPR004009; Myosin\_N.

DR Interpro: IPR001609; myosin\_head.

DR Pfam: PF00612; IQ; 2.

DR Pfam: PF00063; myosin\_head; 1.

DR Pfam: PF02736; myosin\_N; 1.

DR PRINTS: PR00193; MYOSINHEAVY.

DR ProDom: PD000355; myosin\_head; 1.

DR SMART: SM00015; IQ; 1.

DR SMART: SM00242; MYSC; 1.

DR PROSITE: PS50096; IQ; 1.

DR Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding; Methylation; Alkylation; Phosphorylation; Multigene family.

FT DOMAIN 1 789

FT DOMAIN 790 819

FT DOMAIN 848 1509

FT DOMAIN 848 1226

FT DOMAIN 1227 1252

FT DOMAIN 1253 1509

FT DOMAIN 1253 1482

FT DOMAIN 1483 1509

FT NP\_BIND 182 189

FT DOMAIN 660 682

FT DOMAIN 766 780

FT MOD\_RES 133 133

FT MOD\_RES 700 700

FT MOD\_RES 1489 1489

FT MOD\_RES 1494 1494

FT MOD\_RES 1499 1499

SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 14.5%; Score 154.5; DB 1; Length 1509;

Best Local Similarity 25.8%; Pred. No. 0.15; Mismatches 55; Gaps 7;

Matches 63; Conservative 52; Indels 91; Indels 55; Gaps 7;

QY 5 NLFEDLAQEEENVLDREFLNELNDVRAQLSDQ-----REKRDSQ 46

DB 894 DKLEKDLAALKLKILDEGEKADLEEDNALLQKKVAGLEELQETSASNDILPQKRLE 953

QY 47 VIITLADTLERNATVYVSLQALGKAE-----MCSILKQMKMLLEQQD 92

DB 954 AEKGLASLEERNRKALQEAQTVSEENELQDYEDAAAHDSLKKEEDLSREL 1013

QY 93 ETKQAEAGR---LSKMKTMQEIQLLOSQPEV-EEKIRDMGQSGAVQOLAVYCV 147

DB 1014 ETKALADAEISETLSKLNTERGADVNEEDVATATLQLEKTKKSLSEELAQTRA 1073

QY 148 SLKREYENLKEARRASGEVADKLKRLDFSSRSKLQTVYSELQAKLELSAQ---KDLQ- 203

DB 1074 QLEBE---KSGKRAASSKAKOLQOQLEDARSEVLSKLSAERKSLTKAKDQNDRLDE 1129

QY 204 -----SADKEIMSLKKILTMLO 220

DB 1130 QLEDERTVRANDVKQKALEKLELE 1156

RESULT 18

MYH2\_HUMAN STANDARD; PRT; 1941 AA.

AC Q9UKX2; Q16229; Q14322;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain IIe) (MYHc-IIa).

GN MYH2 OR MYHSA2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxId=9606;

ON [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Skeletal muscle;

RX MEDLINE=9931869; PubMed=10388558;

RA Weiss A., Schiaffino S., Leinwand L.A.;

RT "Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: Implications for functional diversity.";

RL J. Mol. Biol. 290:61-75(1999).

RN [2]

RP SEQUENCE OF 1711-1941 FROM N.A.

RC TISSUE=Skeletal muscle;

RX MEDLINE=95109625; PubMed=7545970;

RA Smerdu V., Karsch-Mizrachi I., Campione M., Leinwand L., Schiaffino S.;

RT "Type IIX myosin heavy chain transcripts are expressed in type IIB fibers of human skeletal muscle.";

RL Am. J. Physiol. 267:C1723-C1728(1994).

RN [3]

RP SEQUENCE OF 1823-1941 FROM N.A.

RC TISSUE=Skeletal muscle;

RX MEDLINE=95270723; PubMed=7751403;

RA Emilion S., Sant'ana Pereira J., Sargeant T., Young A., Goldspink G.;

RT "Characterization of human skeletal muscle fibres according to the myosin heavy chains they express.";

RL J. Muscle Res. Cell Motil. 16:35-43(1995).

CC -1- FUNCTION: MUSCLE CONTRACTION.

CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AF111784; AAD29950.1; -

DR EMBL: S73840; AAC13916.1; -

DR EMBL; 232858; CAA83687.1; -  
 DR HSSP; P08799; 1MND.  
 DR MM; 160740; -  
 DR InterPro; IPR000048; IQ.  
 DR InterPro; IPR004009; Myosin\_N.  
 DR InterPro; IPR002928; Myosin\_tail.  
 DR InterPro; IPR001609; myosin\_head.  
 DR Pfam; PF00612; IQ; 2.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PSS0096; IQ; 1.  
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KM Multigene family.  
 FT DOMAIN 1 786 MYOSIN HEAD-LIKE.  
 FT DOMAIN 787 816 IQ.  
 FT DOMAIN 845 1941 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT CONFLICT 1844 1844 K -> R (IN REF. 3).  
 SQ SEQUENCE 1941 AA; 223043 MW; 681E866F83AE83F CRC64;

Query Match 14.4%; Score 154; DB 1; Length 1941;  
 Best Local Similarity 23.3%; Pred. No. 0.21;  
 Matches 53; Conservative 51; Mismatches 79; Indels 44; Gaps 7;

QY 24 LKNEIDNRAQLSQKDKRDSQVIIDTLRDLTEERNATVSLQALGKAEMLSTLKKQ 83  
 Db 1209 LGEQIDNLRKQKLEKSEKMEIDLASNV-----TSKAKNLEKCTLEDQ 1261  
 QY 84 MKYLEQODETK-----QAQEEAGRLE-----RSKAKTMEQIELL 117  
 Db 1262 LSELSKSEEEQRLINDTLTAORGLQTESGEFSRQDEKALVSLSGKQAFQOIEL 1321  
 QY 118 LOSQLPEVEEMIRDMGQSGAVEQLAVYCVSLKKEYLEMKARRASGEVAKLRDLFSS 177  
 Db 1322 KR----QLEELIKAKNALAHLSQSRHDCDLLEQYEEQESKALEQLRALSANTVAQW 1377  
 QY 178 RSKLQIOT---YSELDAQLELKSQKDLQSADEKIMSLKKKLTMLQ 220  
 Db 1378 RTKYTEIDAIQRTLELEAKKLL-AQR-LQAAEHVEAVNAKCSLE 1421

RESULT 19  
 G160\_HUMAN STANDARD; PRT; 579 AA.  
 AC 008378:  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Golgin-160 (Fragment).  
 GN GOLGA3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver.  
 RX MEDLINE=93301617; PubMed=8315394;  
 RA Fitzler M.J., Hamel J.C., Ochs R.L., Chan E.K.L.;  
 RT "Molecular characterization of two human autoantigens: unique CDNA  
 RT encoding 95- and 160-kD proteins of a putative family in the golgi  
 RT complex";  
 RL J. Exp. Med. 178:49-62(1993).  
 CC -1- FUNCTION: GOLGI AUTO-ANTIGEN. MAY HAVE A FUNCTION IN THE  
 CC PROCESSING AND TRANSPORT OF PROTEINS THROUGH THE GOLGI.  
 CC -1- SIMILARITY: HIGH, TO MOUSE MALE-ENHANCED ANTIGEN-2 (MEA-2).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L06148; AAA35921.1; -  
 DR HSSP; P80220; IDIP.  
 DR MM; 602581; -  
 DR Golgi stack.  
 FT NON\_TER 1 1  
 FT NON\_TER 579 579  
 SQ SEQUENCE 579 AA; 66350 MW; 8C9A5A73D0128AEE CRC64;

Query Match 14.3%; Score 152.5; DB 1; Length 579;  
 Best Local Similarity 25.0%; Pred. No. 0.075;  
 Matches 55; Conservative 42; Mismatches 86; Indels 37; Gaps 6;

QY 13 QEEENVLDREFLKNELD-----NVRQLSQKDKRDSQVIIDTLRDLTEERNATVSL 66  
 Db 363 REHNSILETALAKREADVQLNLQAVLQRRKEBEDROMKHLVQALQASLEKEKRYNSL 422  
 QY 67 QQALGKAEMLCSTLKKQKYLEQODETK-----QAQEEAGRLEKSKTME- 112  
 Db 423 KEQYAAAVGEGNRRHRAKASLESEVKELQAKKHLVQKLAQAEADLQIREGKHSQEI 482  
 QY 113 ---QIEL-----LOSQLPEVEEMIRDMGQSGAVEQLAVYCVSLKKEYLEMKARRK 161  
 Db 483 AQQAELAEARQQLQQLQKQ-----DEQLSKQPVNGQEMENLKMEVDQKEREIQLKQQLD 539  
 QY 162 ASGEVADKLKQDLFSSRSKLTQTVYSELDAQLELKSQKDL 201  
 Db 540 LTEQOG---RKLEGLQLOLLQNVKSELEMAQEDLSMTOKD 576

RESULT 20  
 MYH7\_HUMAN STANDARD; PRT; 1935 AA.  
 AC P12883; Q14904; Q16579;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, cardiac muscle beta isoform (MyHC-beta).  
 GN MYH7 OR MYHCB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90301496; PubMed=2249844;  
 RA Jaenike T., Diederich K.W., Haas W., Schleich J., Lichter P.,  
 RA Pfordt M., Bach A., Vosberg H.P.;  
 RT "The complete sequence of the human beta-myosin heavy chain gene and  
 RT a comparative analysis of its product";  
 RL Genomics 8:194-206(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90301496; PubMed=2362820;  
 RA Liew C.-C., Sole M.J., Yamauchi-Takahara K., Kellam B.,  
 RA Anderson D.H., Lin L., Liew J.;  
 RT "Complete sequence and organization of the human cardiac beta-myosin  
 RT heavy chain gene";  
 RL Nucleic Acids Res. 18:3647-3651(1990).  
 RN [3]  
 RP SEQUENCE OF 1-115 FROM N.A.  
 RX MEDLINE=89264452; PubMed=2726733;  
 RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;  
 RT "Characterization of human cardiac myosin heavy chain genes";

- RL Proc. Natl. Acad. Sci. U.S.A. 86:3504-3508(1989).  
 RN [14]  
 RP ERRATUM.  
 RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7416-7417(1989).  
 RN [15]  
 RP SEQUENCE OF 1310-1935 FROM N.A.  
 RX MEDLINE=86176778; PubMed=2421254;  
 RA Saez L., Leinwand L.A.;  
 RT "Characterization of diverse forms of myosin heavy chain expressed in  
 RL adult human skeletal muscle.";  
 RN Nucleic Acids Res. 14:2951-2969(1986).  
 RP REVISIONS.  
 RA Leinwand L.A.;  
 RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.  
 RN [17]  
 RP SEQUENCE OF 1410-1935 FROM N.A.  
 RX MEDLINE=88299163; PubMed=2969919;  
 RA Kurabayashi M., Tsuchimochi H., Komuro I., Takaku F., Yazaki Y.;  
 RT "Molecular cloning and characterization of human cardiac alpha- and  
 RT beta-form myosin heavy chain complementary DNA clones. Regulation of  
 RT expression during development and pressure overload in human  
 RT atrium.";  
 RL J. Clin. Invest. 82:524-531(1988).  
 RN [8]  
 RP SEQUENCE OF 785-1935 FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=90235862; PubMed=1691980;  
 RA Bobber E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,  
 RA Arnold H.H.;  
 RT "Identification of three developmentally controlled isoforms of human  
 RT myosin heavy chains.";  
 RL Eur. J. Biochem. 189:55-65(1990).  
 RN [19]  
 RP SEQUENCE OF 1393-1935 FROM N.A.  
 RX MEDLINE=87192738; PubMed=3032769;  
 RA Jandreski M.A., Liew C.-C.;  
 RT "Construction of a human ventricular cDNA library and  
 RT characterization of a beta myosin heavy chain cDNA clone.";  
 RL Hum. Genet. 76:47-53(1987).  
 RN [10]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=96039076; PubMed=8533830;  
 RA Arai S., Matsuo K., Hirayama K., Sukurai H., Tamura M., Ozawa T.,  
 RA Kimura M., Imanura S.-I., Furutani Y., Joh-O K., Kawana M., Takao A.,  
 RA Hosoda S., Momma K.;  
 RT "Missense mutation of the beta-cardiac myosin heavy-chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Am. J. Med. Genet. 58:267-276(1995).  
 RN [11]  
 RP VARIANTS CMH1 GLU-256 AND ARG-741.  
 RX MEDLINE=93248216; PubMed=8483915;  
 RA Pananapazir L., Dalakas M.C., Cyran F., Cohn G., Epstein N.D.;  
 RT "Missense mutations in the beta-myosin heavy-chain gene cause central  
 RT core disease in hypertrophic cardiomyopathy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3993-3997(1993).  
 RN [12]  
 RP VARIANT CMH1 GLN-403.  
 RX MEDLINE=90367131; PubMed=1975517;  
 RA Geisterfer-Lowrance A.A.T., Kass S., Tanigawa G., Vosberg H.-P.,  
 RA McKenna W., Seidman C.E., Seidman J.G.;  
 RT "A molecular basis for familial hypertrophic cardiomyopathy: a beta  
 RT cardiac myosin heavy chain gene missense mutation.";  
 RL Cell 62:999-1006(1990).  
 RN [13]  
 RP VARIANTS CMH1 GLN-249; GLN-403; CYS-453 AND MET-606.  
 RX MEDLINE=92204193; PubMed=1552912;  
 RA Watkins H., Rosenzweig A., Hwang D.S., Levi T., McKenna W.,  
 RA Seidmann C.E., Seidmann J.G.;  
 RT "Characteristics and prognostic implications of myosin missense  
 RT mutations in familial hypertrophic cardiomyopathy.";  
 RL New Engl. J. Med. 326:1108-1114(1992).
- RN [14]  
 RP VARIANTS CMH1 GLN-403; CYS-453; ARG-584 AND MET-606.  
 RX MEDLINE=94070863; PubMed=8250038;  
 RA Watkins H., Thierfelder L., Anan R., Jarcho J., Matsumori A.,  
 RA McKenna W., Seidman J.G., Seidman C.E.;  
 RT "Independent origin of identical beta cardiac myosin heavy-chain  
 RT mutations in hypertrophic cardiomyopathy.";  
 RL Am. J. Hum. Genet. 53:1180-1185(1993).  
 RN [15]  
 RP VARIANTS CMH1 GLN-403 AND VAL-908.  
 RX MEDLINE=92346810; PubMed=1638703;  
 RA Epstein N.D., Cohn G.M., Cyran F., Pananapazir L.;  
 RT "Differences in clinical expression of hypertrophic cardiomyopathy  
 RT associated with two distinct mutations in the beta-myosin heavy chain  
 RT gene. A 908Leu-->Val mutation and a 403Arg-->Gln mutation.";  
 RL Circulation 86:345-352(1992).  
 RN [16]  
 RP VARIANTS CMH1 LEU-403 AND TRP-403.  
 RX MEDLINE=94075629; PubMed=8254035;  
 RA Dausse E., Komajda M., Feller L., Dubourg O., Dufour C., Carrier L.,  
 RA Wanssely C., Bericovic J., Hengstenberg C., Al-Mahdawi S.;  
 RT "Familial hypertrophic cardiomyopathy. Microsatellite haplotyping and  
 RT identification of a hot spot for mutations in the beta-myosin heavy  
 RT chain gene.";  
 RL J. Clin. Invest. 92:2807-2813(1993).  
 RN [17]  
 RP VARIANT CMH1 TRP-403.  
 RX MEDLINE=94093568; PubMed=8268932;  
 RA Woolman J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a new missense mutation at Arg403, a Cpg mutation  
 RT hotspot, in exon 13 of the beta-myosin heavy chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 2:1731-1732(1993).  
 RN [18]  
 RP VARIANT CMH1 ASN-615.  
 RX MEDLINE=93038688; PubMed=1417858;  
 RA Nishi H., Kimura A., Harada H., Yoshima H., Sasazuki T.;  
 RT "Novel missense mutation in cardiac beta myosin heavy chain gene  
 RT found in a Japanese patient with hypertrophic cardiomyopathy.";  
 RL Biochem. Biophys. Res. Commun. 188:379-387(1992).  
 RN [19]  
 RP VARIANT CMH1 GLY-778.  
 RX MEDLINE=93343938; PubMed=8343162;  
 RA Harada H., Kimura A., Nishi H., Sasazuki T., Yoshima H.;  
 RT "A missense mutation of cardiac beta-myosin heavy chain gene linked  
 RT to familial hypertrophic cardiomyopathy in affected Japanese  
 RT families.";  
 RL Biochem. Biophys. Res. Commun. 194:791-798(1993).  
 RN [20]  
 RP VARIANT CMH1 VAL-908.  
 RX MEDLINE=93168485; PubMed=8435239;  
 RA Al-Mahdawi S., Chamberlain S., Cleland J., Nihoyannopoulos P.,  
 RA Gilligan D., French J., Choudhury L., Williamson R., Oakley C.;  
 RT "Identification of a mutation in the beta cardiac myosin heavy chain  
 RT gene in a family with hypertrophic cardiomyopathy.";  
 RL Br. Heart J. 69:136-141(1993).  
 RN [21]  
 RP VARIANT CMH1 TRP-719.  
 RX MEDLINE=95179132; PubMed=7874131;  
 RA Greve G., Bachinski L., Friedman D.L., Czernezewicz G., Anan R.,  
 RA Towbin J., Seidman C.E., Roberts R.;  
 RT "Isolation of a de novo mutant myocardial beta MHC protein in a  
 RT pedigree with hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 3:2073-2075(1994).  
 RN [22]  
 RP VARIANTS CMH1 CYS-513; ARG-716 AND TRP-719.  
 RX MEDLINE=94110336; PubMed=8282798;  
 RA Anan R., Greve G., Thierfelder L., Watkins H., McKenna W., Solomon S.,  
 RA Vecchio C., Shono H., Nakao S., Tanaka H., Mares A. Jr., Towbin J.A.,  
 RA Spirito P., Roberts R., Seidman J.G., Seidman C.E.;  
 RT "Prognostic implications of novel beta cardiac myosin heavy chain gene  
 RT mutations that cause familial hypertrophic cardiomyopathy.";  
 RL J. Clin. Invest. 93:280-285(1994).

RN [23]  
 RP VARIANT CMH1 THR-797.  
 RX MEDLINE=96047159; PubMed=7581410;  
 RA Moolman J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a novel Ala797Thr mutation in exon 21 of the beta-  
 RL myosin heavy chain gene in hypertrophic cardiomyopathy.";  
 RN Hum. Mutat. 6:197-198(1995).  
 RN [24]  
 RP VARIANT CMH1 CYS-453.  
 RX MEDLINE=96209901; PubMed=8655135;  
 RA Ko Y.-L., Chen J.-J., Tang T.-K., Cheng J.-J., Lin S.-Y., Liou Y.-C.,  
 RA Kuan P., Wu C.-W., Lien W.-P., Liew C.-C.;  
 RT "Malignant familial hypertrophic cardiomyopathy in a family with a  
 RT 453Arg->Cys mutation in the beta-myosin heavy chain gene:  
 RT coexistence of sudden death and end-stage heart failure.";  
 RL Hum. Genet. 97:585-590(1996).  
 RN [25]  
 RP VARIANTS CMH1 THR-349 AND TRP-719.  
 RX MEDLINE=98204402; PubMed=9544842;  
 RA Jesche B., Uhl K., Weist B., Schröder D., Meltinger T.,  
 RA Dohlemann C., Vosberg H.-P.;  
 RT "A high risk phenotype of hypertrophic cardiomyopathy associated with  
 RT Query Match 14.3%; Score 152.5; DB 1; Length 1935;  
 Best Local Similarity 20.5%; Pred. No. 0.25;  
 Matches 61; Conservative 60; Mismatches 88; Indels 89; Gaps 9;  
 QY 1 RTIINKLFFDLAEEENLDR-----EFLK----- 25  
 Db 1129 RAKVEKLSDSLRELESEERLEEGATSVQIEMKKREAFQMRDLEATLQHEAT 1188  
 QY 26 -----NELDNVRAQLSQDKREKDSQVYIITLRLTEERNATVYLSQA 69  
 Db 1189 AALRLKKHADSVAELGEGIDNLQRYKQKLEKSEFKLEDDVYSNMQ-----IKKA 1241  
 QY 70 LGAEMLCSTLKKQMKYLEDQODETKQAQOEAGRLSRMKT-----MEQIELLSQ 121  
 Db 1242 KANLEKMCRTLEDQNEHRSKAEETORSVNDLTSQRAKLQENGLSHQDEKELI-SQ 1300  
 QY 122 LP-----EVEEMIRMGVGSADVEDLAVYCVSLKKEVENLLEAKRASEV 166  
 Db 1301 LTRGKLYTQOLEDLKLROLEBEVAKKNAHLAHLQSRHDCDLRQYEEETEAKAELOVR 1360  
 QY 167 ADLRLKDLFSSRSKLOT-----VSELDQAKLELSAQKDLQADKEISLKKITLMQ 220  
 Db 1361 LSKANSEVAQMRKTEYDAIQRTLELEAKKL--AQR-LQEAEEAVAVNAKCSLE 1415  
 RESULT 21  
 MYSA\_DROME STANDARD: PRT; 2411 AA.  
 AC POS661:  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, muscle.  
 GN MHC.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89384536; PubMed=2506434;  
 RA George E.L., Ober M.B., Emerson C.P. Jr.;  
 RT "Functional domains of the Drosophila melanogaster muscle myosin  
 RT heavy-chain gene are encoded by alternatively spliced exons.";  
 RL Mol. Cell. Biol. 9:2957-2974(1989).  
 RN [2]  
 RP SEQUENCE OF 1-312 FROM N.A.  
 RX MEDLINE=87280141; PubMed=3038896;  
 RA Wassenberg D.R. II, Kronert W.A., O'Donnell P.T., Bernstein S.I.;

RT "Analysis of the 5' end of the Drosophila muscle myosin heavy chain  
 RT gene. Alternatively spliced transcripts initiate at a single site and  
 RT intron locations are conserved compared to myosin genes of other  
 RT organisms.";  
 RL J. Biol. Chem. 262:10741-10747(1987).  
 RN [3]  
 RP SEQUENCE OF 486-881 FROM N.A.  
 RC STRAIN-CANTON-S; TISSUE=Embryonic muscle;  
 RX MEDLINE=91330870; PubMed=1907912;  
 RA Kronert W.A., Edwards K.A., Roche E.S., Wells L., Bernstein S.I.;  
 RT "Muscle-specific accumulation of Drosophila myosin heavy chains: a  
 RT splicing mutation in an alternative exon results in an isoform  
 RT substitution.";  
 RL EMBO J. 10:2479-2488(1991).  
 CC -I- FUNCTION: MUSCLE CONTRACTION.  
 CC -I- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -I- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -I- ALTERNATIVE PRODUCTS: AS DROSOPHILA HAS A SINGLE MUSCLE MHC GENE,  
 CC MHC ISOFORMS MUST RELY ON ALTERNATIVE RNA SPLICING THAT LEADS TO  
 CC DIFFERENCES IN THE C-TERMINUS OF THE VARIOUS MHC PROTEINS.  
 CC -I- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -I- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M61229; AAA28686.1; ALT\_SEQ.  
 DR EMBL: M61229; AAA28687.1; ALT\_SEQ.  
 DR EMBL: J02788; AAA28706.1; ALT\_SEQ.  
 DR EMBL: J02788; AAA28707.1; ALT\_SEQ.  
 DR EMBL: X60196; CAA42752.1; ALT\_SEQ.  
 DR EMBL: X60196; CAA42753.1; ALT\_SEQ.  
 DR EMBL: X60196; CAA42754.1; ALT\_SEQ.  
 DR PIR: A28492; A28492.  
 DR PIR: A32491; A32491.  
 DR PIR: B32491; B32491.  
 DR HSP: P08799; LMND.  
 DR FLYBase: FBgn0002741; Mhc.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 2.  
 DR Pfam: PF00063; myosin\_head; 7.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR PRODom: PD000335; myosin\_head; 5.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 KW Myosin: Muscle protein. Coiled coil. Thick filament. Actin-binding;  
 KW ATP-binding; Calmodulin-binding; Alternative splicing;  
 KW Multigene family.  
 FT DOMAIN 1 ? MYOSIN HEAD-LIKE.  
 FT DOMAIN 1205 1232 IO.  
 FT DOMAIN 1233 2376 COILED COIL (POTENTIAL).  
 FT NP\_BIND 227 234 ATP (BY SIMILARITY).  
 FT VARSPIC 2385 2385 P -> I (IN SHORT ISOFORM).  
 FT VARSPIC 2386 2411 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 43 44 EK -> RE (IN REF. 2).  
 SQ SEQUENCE 2411 AA; 276434 MW; 8C4FEC1BB8D310A9 CRC64;





DR InterPro: IPR002928; Myosin\_tail.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00612; IQ: 1.  
DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF02735; myosin\_N; 1.  
DR Pfam: PF01576; Myosin\_tail; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR PRODOM: PD000355; myosin\_head; 1.  
DR SMART: SM00015; IQ: 1.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS50096; IQ: 1.  
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Methylation; Alkylation; Multigene family.  
FT DOMAIN 1 793  
FT DOMAIN 794 823  
FT NP\_BIND 857 1969  
FT NP\_BIND 179 186  
FT DOMAIN 667 689  
FT DOMAIN 770 784  
FT MOD\_RES 130 130  
FT MOD\_RES 707 707  
FT MOD\_RES 717 717  
SO SEQUENCE 1969 AA; 225509 MW; 64577BBA7EAD80A CRC64;

Query Match 14.1%; Score 150; DB 1; Length 1969;  
Best Local Similarity 23.5%; Pred. No. 0.35;  
Matches 64; Conservative 48; Mismatches 94; Indels 66; Gaps 10;

QY 5 NKLFPDLAQBENYLDREFLNELNDVRAQLSQK-----DKEKRD 45  
DB 898 NALFSLTEKCANLADAEERNEKLNQLKATLESKLSIDITGQLEDWQERNEDLARKKKTID 957

QY 46 QVIIDT-----LBDTLEE---RNATVVSQQALGKAMCSTLKKQKTYEPOOD 92  
DB 958 QELSTQKKNVODLESLRKABQEKSRDHNIRSLQDENANODEAVAKLNKEKKH---OEE 1014

QY 93 ETQAOEAGRLSKMTMEQIELLQSLPEVEEMI---RDMVGQSAVEQLAVYCVS 148  
DB 1015 SNRKINDELQSEEDKVNLEKIRNKLQMDLEENIDREKRSRDIKARKKVE---GD 1071

QY 149 LKKEYNLEKARKASGEVADKLK---DLFSSR-----SKLQTVYSEL----- 188  
DB 1072 LKVAQENIDEITKQKHVDYETLLKREEDLHHTNAKLAEINNSIAKLQSLIKELFARNAL 1131

QY 189 -DQAKLEKSNQKXDS---ADKEIMSLKKTL 216  
DB 1132 EEELEARNRSKSDRSRSEARELELETERL 1163

RESULT 24  
KUCR\_MOUSE STANDARD: PRT: 548 AA.  
AC P70194;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Kuppfer cell receptor.  
CN KCLR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=Liver;  
RA Takekawa R., Magatsuma H., Nomoto C., Watanabe Y., Akaike T.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.  
CC COULD BE INVOLVED IN ENDOCYTOSIS.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -1- TISSUE SPECIFICITY: KUPFER CELLS.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC -----  
CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
CC -----

DR EMBL: D88577; BA113647.1; -.  
DR HSSP: P20693; HLJ.  
DR MGD: MGI:1859834; KCLR.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00059; lectin\_c; 1.  
DR SMART: SM00034; CLECT; 1.  
DR PROSITE: PS00615; C-TYPE\_LLECTIN\_1; 1.  
DR PROSITE: PS50041; C-TYPE\_LLECTIN\_2; 1.  
KW Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;  
KW Endocytosis.  
FT DOMAIN 1 42  
FT TRANSMEM 43 69  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT DOMAIN 438 548  
FT DOMAIN 438 538  
FT DISULFID 440 536  
FT DISULFID 516 528  
FT CARBOHYD 86 86  
FT CARBOHYD 92 92  
FT CARBOHYD 115 115  
FT CARBOHYD 132 132  
FT CARBOHYD 209 209  
FT CARBOHYD 255 255  
SO SEQUENCE 548 AA; 61268 MW; 6F6495E820E73BD9 CRC64;

Query Match 14.0%; Score 149; DB 1; Length 548;  
Best Local Similarity 26.0%; Pred. No. 0.11;  
Matches 59; Conservative 45; Mismatches 85; Indels 38; Gaps 9;

QY 22 EFLKNELDNVAQLSQ-KDREKRSQVIIDT-LRDTLEERNATVVSQQALGKAMC 78  
DB 138 QLSHLELDVAVADILQTDVLEKSGALALETQALRSLEVASADHSIRGDELEKANAMTS 197

QY 79 TLKQMKY-----LEQQODETKQAQEGRLRSKKKMTMEQIELLQSLP 123  
DB 198 QTRGLKSTENTSALHVLGRGLEAQS---IOALRGLQASANDLSSQTGFLQHSMD 254

QY 124 EEEIMRDMVGQSAVEQLAVYCVSLKREYENTL-KEARKASGEVADK-----LRKDLFS 176  
DB 255 NISAQIQTV---RDMERAGEKMSLKELETLTAQTOKANGHLEQTDQIQGLKAEKLS 311

QY 177 S---RSKLQTVYSELDOAKLEKSNQKXDSADKEIMSLKKTLMLQ 220  
DB 312 TSSUNSLRIEYVVGQMKDASRELQTLRRDL---SDVSALKSNVOMLQ 354

RESULT 25  
RAS0\_AERPE STANDARD: PRT: 919 AA.  
AC Q9YFZ1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DNA double-strand break repair rad50 ATPase.  
GN RAD50 OR APE0110.  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;  
OX Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KI;  
RX MEDLINE=99310339; PubMed=10382966;



FT	CONFLICT	1473	1474	AS -> SF (IN REF. 2).
FT	CONFLICT	1569	1569	L -> V (IN REF. 2).
FT	CONFLICT	1598	1598	V -> E (IN REF. 2).
FT	CONFLICT	1606	1606	D -> N (IN REF. 2).
FT	CONFLICT	1643	1643	A -> D (IN REF. 2).
FT	CONFLICT	1648	1648	R -> O (IN REF. 2).
FT	CONFLICT	1750	1750	O -> K (IN REF. 2).
FT	CONFLICT	1822	1822	R -> K (IN REF. 2).
FT	CONFLICT	1845	1845	R -> H (IN REF. 2).
SO	SEQUENCE	1939	AA; 223114	MW; 39ADB26AB79DFA53 CRC64;

  

Query Match	14.0%	Score 149:	DB 1:	Length 1939:
Best Local Similarity	22.9%	Pred. No. 0.39:		
Matches 52:	Conservative	50:	Mismatches 81:	Indels 44:
				Gaps 7:

  

QY	24	LKNELDNVRALDSOKDKKRDSDOVIITDRLPTLEFRNATVVSLOALCAKMLCSTLTKQ	83
Db	1207	LGEQIDNLRQKQKRLERKSEKMEKIDDLASNM-----TVSAKKNLEKMCRALED	1259
QY	84	MKYLEQODETR-----QAQEBAGRL-----RSKKTKYQIEL	117
Db	1260	LSEITKKEEQORLINDTAAORARLQTESGESRQDLEKDTLVLSQSGKQAFQIQIEEL	1319
QY	118	LQSLPEVEEMIRDMGVQSAVEQOLAVYCVSLKKEKYLEKARKASGEVADKRLDLESS	177
Db	1320	KR-----QLEELIKAKSALAHALDSSRHCDLLRDEYEEQEKELQRAMKANSEVQW	1375
QY	178	RSKLTQ-----VYSELQAKLELKSQKDLQSDAKDEIMSLKRLMLQ	220
Db	1376	RTKYETDAILQRTLEELKAKKRL--AQ-RLQAEHVEAVNKKASLE	1419

  

RESULT 28
RA50_ARCFU STANDARD: PRT: 866 AA.
ID RA50_ARCFU STANDARD: PRT: 866 AA.
AC 029230:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA double-strand break repair rad50 Arpase.
GN RAD50 OR AFI1032.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
XX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=96049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham M.D., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirtness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glöckner A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Usterback T.,
RA Coston M.D., Spriggs T., Artchak P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -!- SUBUNIT: Forms a complex with mre11 (by similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.



```

OY 168 DKLRDLFSSR-----SKLOTVYSELDQA-----KLELKSQKDLQSGADKEI 209
DB 665 DALSEELVOLRAQEKVHEMEKEHLNKVQTA-NEVKAQAVEQIQSHRETHQKQISSLRDEV 723
OY 210 MSKKKLTMLQ 220
DB 724 EAKAKLITDLQ 734

RESULT 30
MYH7_MESAU STANDARD: PRT: 1934 AA.
AC P13540; Q60540;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, cardiac muscle beta isoform (myHC-beta).
GN MYH7.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FLB; TISSUE=Liver;
RX MEDLINE=95115033; PubMed=7815459;
RA Wang R., Sole M.J., Cukerman E., Liew C.-C.;
RT "Characterization and nucleotide sequence of the cardiac alpha-myosin
RT heavy chain gene from Syrian hamster.";
RL J. Mol. Cell. Cardiol. 26:1155-1165(1994).
RN [2]
RP SEQUENCE OF 962-1935 FROM N.A.
RX MEDLINE=88247788; PubMed=3380703;
RA Jandreski M.A., Sole M.J., Liew C.-C.;
RT "Sequence of cDNA encoding the Syrian hamster cardiac beta-myosin
RT heavy chain.";
RL Nucleic Acids Res. 16:4737-4737(1988).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (SL) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L12104; AAA62313.1; -
DR EMBL; X07273; CAA30256.1; -
DR PIR; A28298; A28298.
DR HSSP; P08799; IAMD.
DR InterPro; IPR000048; IO.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head.
DR Pfam; PF02736; myosin_N; 1.
DR Pfam; PF01576; myosin_tail; 1.

```

```

DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family;
KW Calmodulin-binding.
FT DOMAIN 1 779 MYOSIN HEAD-LIKE.
FT DOMAIN 780 809 IO.
FT DOMAIN 839 1934 COILED COIL (POTENTIAL).
FT NP_BIND 177 184 ATP.
FT DOMAIN 654 676 ACTIN-BINDING.
FT DOMAIN 756 770 ACTIN-BINDING.
FT MOD_RES 128 128 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 694 694 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
FT MOD_RES 966 966 D->E (IN REF. 2).
FT CONFLICT 966 966 T->TE (IN REF. 2).
FT CONFLICT 978 978 E->Q (IN REF. 2).
FT CONFLICT 986 986 DLOAED -> ALBARKT (IN REF. 2).
FT CONFLICT 1008 1014 D->Y (IN REF. 2).
FT CONFLICT 1057 1057 L->V (IN REF. 2).
FT CONFLICT 1060 1060 D->N (IN REF. 2).
FT CONFLICT 1095 1095 E->D (IN REF. 2).
FT CONFLICT 1217 1217 D->N (IN REF. 2).
FT CONFLICT 1271 1271 T->A (IN REF. 2).
FT CONFLICT 1327 1327 C->R (IN REF. 2).
FT CONFLICT 1358 1358 L->V (IN REF. 2).
FT CONFLICT 1504 1504 M->L (IN REF. 2).
FT CONFLICT 1537 1537 N->K (IN REF. 2).
FT CONFLICT 1556 1556
SQ SEQUENCE 1934 AA; 222928 MW; FDBAG58310B0B57D CRC64;

```

```

Query Match 13.9%; Score 148.5; DB 1; Length 1934;
Best Local Similarity 20.5%; Pred. No. 0.42;
Matches 61; Conservative 60; Mismatches 88; Indels 89; Gaps 9;

```

```

OY 1 RTIKLFFPDIAEEENVLDL-----EFLK----- 25
DB 1128 RAKVEKLSDLSRELEETSEKLEFAGGATSVQIEANKKREAFQKMRDLEATLQHEAT 1187
OY 26 -----NELDNVRAQLSQDKREKRDQVYITDLRDTLEARNATVVSLOQA 69
DB 1188 AAALRRKHADSVAVELGEQIDNLQRYKQLEKEKSEFKLELDVYSNMQ-----IKA 1240
OY 70 LGRKEMLCSTLKKQKTYLEQOODETKQAEAGRLRSMTK-----MQEILLDSQ 121
DB 1241 KANLEKMCRTLEDQNEHRSKAEETQRSVNDLTSGRAKLTQENGLSRQDKEKALI-SQ 1299
OY 122 LP-----EVEEMIRDMGVGQSAVEQLAVYCVSLKKEYENLKEARKASGEV 166
DB 1300 LTRGKLTYYTOLEDLKRLEEEVAKAKNTLAHLASARDCDLRLQRYEEETFAKAELOCV 1359
OY 167 ADLKRDLFSSRSKLOT-----VYSELDQAKLELSAQKRDQADKEIMSLKKLTMLQ 220
DB 1360 LSKANSEVAQQRTRYETDAIQRTLELEAKKKL--AQK-LQDAEBAVAVNAKSSLE 1414

RESULT 31
MYSS_CHICK STANDARD: PRT: 1938 AA.
AC P13538; O13228;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

```

RP SEQUENCE FROM N.A.  
 RC STRAIN-WHITE LEGHORN; TISSUE=Pectoralis muscle;  
 RA Chao T.H., Bandman E., Moore L.;  
 RT "Cloning, nucleotide sequence and characterization of a full-length  
 RT cDNA encoding the myosin heavy chain from adult chicken pectoralis  
 RT major muscle.";  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [12]  
 RP SEQUENCE OF 1-205.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041767; PubMed=1939027;  
 RA Hayashida M., Maiba T., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: I.  
 RT Sequence of the amino-terminal 23 kDa fragment.";  
 RL J. Biochem. 110:54-59(1991).  
 RN [3]  
 RP SEQUENCE OF 206-636.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041768; PubMed=1939028;  
 RA Komine Y., Maiba T., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: II.  
 RT Sequence of the 50 kDa fragment of subfragment-1.";  
 RL J. Biochem. 110:60-67(1991).  
 RN [4]  
 RP SEQUENCE OF 637-837.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041769; PubMed=1939029;  
 RA Maiba T., Miyaniishi T., Matsuzono K., Tanioaka Y., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: III.  
 RT Sequence of the 22 kDa fragment and the alignment of the 23 kDa, 50  
 RT kDa, and 22 kDa fragments.";  
 RL J. Biochem. 110:68-74(1991).  
 RN [5]  
 RP SEQUENCE OF 838-1938.  
 RC TISSUE=Pectoralis muscle;  
 RX MEDLINE=92041770; PubMed=1939030;  
 RA Maiba T., Yajima E., Nagata S., Miyaniishi T., Nakayama S., Matsuda G.;  
 RT "The primary structure of skeletal muscle myosin heavy chain: IV.  
 RT Sequence of the rod, and the complete 1,938-residue sequence of the  
 RT heavy chain.";  
 RL J. Biochem. 110:75-87(1991).  
 RN [6]  
 RP PRELIMINARY SEQUENCE OF 1-808.  
 RX MEDLINE=87092420; PubMed=3467365;  
 RA Maiba T., Hayashida M., Tanioaka Y., Komine Y., Matsuda G.;  
 RT "The primary structure of the myosin head.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:416-420(1987).  
 RN [7]  
 RP SEQUENCE OF 842-1270.  
 RX MEDLINE=90121764; PubMed=2610940;  
 RA Matanabe B.;  
 RT "Complete amino-acid sequence of subfragment-2 in adult chicken  
 RT skeletal muscle myosin.";  
 RL Biol. Chem. Hoppe-Seyler 370:1027-1034(1989).  
 RN [8]  
 RP SEQUENCE OF 852-1108.  
 RX MEDLINE=89374803; PubMed=2775482;  
 RA Matanabe B.;  
 RT "Amino-acid sequence of the short subfragment-2 in adult chicken  
 RT skeletal muscle myosin.";  
 RL Biol. Chem. Hoppe-Seyler 370:549-558(1989).  
 RN [9]  
 RP SEQUENCE OF 1145-1270.  
 RX MEDLINE=89228549; PubMed=2713098;  
 RA Matanabe B.;  
 RT "Amino-acid sequence of the hinge region in chicken myosin  
 RT subfragment-2.";  
 RL Biol. Chem. Hoppe-Seyler 370:55-61(1989).  
 RN [10]  
 RP SEQUENCE OF 1857-1938 FROM N.A.  
 RX MEDLINE=87217964; PubMed=3034534;  
 RA Moriarty D.M., Barringer K.J., Dodgson J.B., Richter H.E.,  
 RA Young R.B.;

RT "Genomic clones encoding chicken myosin heavy-chain genes.";  
 RL DNA 6:91-99(1987).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-843.  
 RX MEDLINE=93303624; PubMed=8316857;  
 RA Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R.,  
 RA Tomchick D.R., Bening M.M., Winkelmann D.A., Wesenberg G.,  
 RA Holden H.M.;  
 RT "Three-dimensional structure of myosin subfragment-1: a molecular  
 RT motor.";  
 RL Science 261:50-58(1993).  
 CC -1- FUNCTION: MUSCLE CONTRACTION. MYOSIN IS A PROTEIN THAT BINDS TO  
 CC F-ACTIN AND HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U87231; AAA47555.1; -;  
 DR EMBL: M16557; AAA48970.1; -;  
 DR PIR: PX0050; PX0051.  
 DR PIR: A26821; A26821.  
 DR PIR: S02082; S02082.  
 DR PIR: S04501; S04501.  
 DR PIR: S05515; S05515.  
 DR PDB: 2MYS; 11-JAN-97.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; Myosin\_head.  
 DR Pfam: PF00612; IQ; 2.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PSS0096; IQ; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;  
 KW Calmodulin-binding; Multigene family; 3D-structure.  
 FT INT\_MET 0  
 FT DOMAIN 1 782 812 MYOSIN HEAD-LIKE.  
 FT DOMAIN 783 838 IO.  
 FT DOMAIN 838 840 HINGE.  
 FT DOMAIN 841 1938 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT DOMAIN 657 679 ACTIN-BINDING.  
 FT DOMAIN 773 773 ACTIN-BINDING.  
 FT MOD\_RES 1 1 ACETYLATION.  
 FT MOD\_RES 35 35 METHYLATION (MONO-).  
 FT MOD\_RES 130 130 METHYLATION (TRI-).  
 FT MOD\_RES 551 551 METHYLATION (TRI-).  
 FT MOD\_RES 755 755 METHYLATION (MONO-).  
 FT

```

FT MOD_RES 697 697 ALKYLATION (SH-1).
FT MOD_RES 707 707 ALKYLATION (SH-2).
FT CONFLICT 907 907 C -> Q (IN REF. 7 AND 8).
FT CONFLICT 980 980 L -> F (IN REF. 1).
FT CONFLICT 1343 1343 E -> D (IN REF. 5).
FT CONFLICT 1545 1545 S -> A (IN REF. 5).
FT CONFLICT 1796 1796 HV -> QL (IN REF. 5).
FT CONFLICT 1830 1830 S -> A (IN REF. 5).
FT CONFLICT 1863 1863 I -> V (IN REF. 10).
FT CONFLICT 1929 1931 IHG -> FH (IN REF. 10).
SO SEQUENCE 1938 AA; 223013 MW; EDD01CBA2681E10F CRC64;

Query Match 13.9%; Score 148.5; DB 1; Length 1938;
Best Local Similarity 22.6%; Pred. No. 0.42;
Matches 53; Conservative 50; Mismatches 80; Indels 51; Gaps 7;

QY 24 LKNELDNRAOLSOQKDKRDSQVYIIDLRTLEERNATVYSLOQALKAEMLCSTLKKQ 83
Db 1205 LGEQIDNLRVKKQKLEKSELMKMDLASNME-----SVSKAKANLEKMCRTLEDQ 1257
QY 84 MKYLEQODE-----TKQAEAGRL-----RSKMTMDEIEL 117
Db 1258 LSEITKEEQNRMINDTQARLQETGYSKQAEKDALISLQSKGQFTQOIEEL 1317
QY 118 LQSLQPEVEEMIRDMGVQSAVEQLAVYCVSLKKEYENLKEAR-----KASGEVAD- 168
Db 1318 KR---HLEEEIKAKNALHALQSAHRDCELLREQYEEOGEAKGELQALSKANSEVAMQW 1373
QY 169 --KLKRDLFSSRSKLQTYVSELDQAKLEKSAQKDLSADKEINSLKRLMLQ 220
Db 1374 RTKYTEIDAIOPTKELEAKRLAQ---RLQDAEEHVEAVNAKCSLEKTKORLQ 1424

RESULT 32
MYH9_CHICK STANDARD; PRT; 1959 AA.
AC P14105;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle (Cellular myosin heavy chain) (NMHC).
GN MYH9.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestinal epithelium;
RX MEDLINE=90046668; PubMed=2813355;
RA Shohet R.V., Conti M.A., Kawamoto S., Preston Y.A., Brill D.A.,
RA Adelstein R.S.;
RT "Cloning of the cDNA encoding the myosin heavy chain of a vertebrate
RT cellular myosin.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7726-7730(1989).
CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPING.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26510; AAA48974.1; -.
CC PIR: A33977; A33977.
CC HSSP: P10587; IBR1.
CC InterPro: IPR000048; IO.
CC InterPro: IPR004009; Myosin_N.
CC InterPro: IPR002928; Myosin_tail.
CC InterPro: IPR001609; myosin_head.
CC Pfam: PF00612; IO; 1.
CC Pfam: PF00063; myosin_head; 1.
CC Pfam: PF02736; Myosin_N; 1.
CC Pfam: PF01576; Myosin_tail; 1.
CC PRINTS: PR00193; MYOSINHEAVY.
CC PRODOM: PD000355; myosin_head; 1.
CC SMART: SM00015; IO; 1.
CC SMART: SM00242; MYSC; 1.
CC PROSITE: PS50096; IO; 1.
CC MYOSIN; ATP-binding; Calmodulin-binding; Actin-binding;
CC Coiled coil; Alkylation; Multigene family.
CC DOMAIN 1 778
CC MYOSIN HEAD-LIKE.
CC DOMAIN 779 808
CC COILED COIL (POTENTIAL).
CC NP_BIND 837 1925
CC ATP (POTENTIAL).
CC DOMAIN 174 181
CC ACTIN-BINDING.
CC MOD_RES 654 676
CC ALKYLATION (SH-1) (POTENTIAL).
CC MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
CC SEQUENCE 1959 AA; 226502 MW; A75C86086FD3A1A1 CRC64;

```

```

Query Match 13.9%; Score 148.5; DB 1; Length 1959;
Best Local Similarity 22.7%; Pred. No. 0.42;
Matches 53; Conservative 56; Mismatches 95; Indels 29; Gaps 8;

QY 11 LAQEEENVLDREFIKNELDNYRAOLSOQKDKRDSQVYIIDLND-----TLEER 59
Db 1665 LAQAKENKRLKSMQAEWIOLOEFLAAERAKROAQOQERDELAINSSGKALAMEK 1724
QY 60 ---NATVYSLOQAL-----GRAEMLCSTLKKQMKYLBQOQDITKQAEAGLRSKMTME 112
Db 1725 RLRLARINQLEELKEEGNTEIINDRIKKNLQIDQNNADLNERNNAQKNENARQOE 1784
QY 113 QTELLSOLPEVEEMIRDMGVQSAVEQLAVYCVSLKKEYE-NLKPARKASGEV--ADK 169
Db 1785 RQNELKIKLQEMESAYKSK--YKATITALEAKTYQLEBQDMDTKERQASKQVRAEK 1842
QY 170 LRKDLF---SSRSKLQTYVSELDQAKLEKSAQKDLSADKEIM--SLKKL 216
Db 1843 KLKIDILLQVDERRNNAQFKQADKAMWRLKQLKRLQEEAEAEQANVRRKL 1895

RESULT 33
MYSB_CAEEL STANDARD; PRT; 1966 AA.
ID MYSB_CAEEL
AC P02566;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain B (MHC B).
GN UNC-54 OR MYO-4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloiderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83273600; PubMed=6576334;
RA Karn J., Brenner S., Barnett L.;
RA "Protein structural domains in the Caenorhabditis elegans unc-54
RT myosin heavy chain gene are not separated by introns.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).

```



```

RN RP SEQUENCE OF 850-1966 FROM N.A.
RX RX MEDLINE=82272395; PubMed=7202124;
RA RA McEachan A.D., Karn J.;
RT RT "Periodic charge distributions in the myosin rod amino acid sequence
RL RL match cross-bridge spacings in muscle.";
RN RN Nature 299:226-231(1982).
[3]
RN RP SEQUENCE OF 1876-1966 FROM N.A.
RX RX MEDLINE=83232892; PubMed=6571695;
RA RA Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
RA RA Waterston R.H.;
RT RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber
RL RL nonsense mutations via altered transfer RNA.";
RN RN Cell 33:575-583(1983).
CC CC -1- FUNCTION: MUSCLE CONTRACTION.
CC CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC CC -1- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC CC -1- PPM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC CC MEMOVOSIN (LMV) AND 1 HEAVY MEMOVOSIN (HMM). IT CAN LATER BE
CC CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC CC SUBFRAGMENT (S2).
CC CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
CC CC C.ELEGANS.
CC CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL; J01050; AAA28124.1; -;
DR DR EMBL; V01494; CAA24738.1; -;
DR DR PIR: A02992; MMYW.
DR DR HSSP; P08799; 1AMD.
DR DR InterPro; IPR004009; Myosin_N.
DR DR InterPro; IPR002928; Myosin_tail.
DR DR InterPro; IPR001609; myosin_head.
DR DR Pfam; PF00063; myosin_head.1.
DR DR Pfam; PF02736; myosin_N.1.
DR DR Pfam; PF01576; Myosin_tail.1.
DR DR PRINTS; PR00193; MYOSINHEAVY.
DR DR PRODOM; PD000355; myosin_head.1.
DR DR SMART; SM00242; MYSC; 1.
KW KW Myosin; muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family.
FT FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
FT FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
FT FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (S2).
FT FT DOMAIN 1165 1176 HINGE.
FT FT DOMAIN 1165 1966 LIGHT MEMOVOSIN (LMV).
FT FT NP_BIND 177 184 ATP (BY SIMILARITY).
FT FT DOMAIN 665 687 ACTIN-BINDING.
FT FT DOMAIN 769 783 ACTIN-BINDING.
FT FT MOD_RES 128 128 METHYLATION (TRI-) (POTENTIAL).
FT FT MOD_RES 705 705 ALKYLATION (SH-1).
FT FT MOD_RES 715 715 ALKYLATION (SH-2).
FT FT CONFLICT 1337 1337 E -> R (IN REF. 2).
FT FT CONFLICT 1880 1880 I -> L (IN REF. 2).
FO FO SEQUENCE 1966 AA; 225125 MW; B66F0B52FE27B67F CRC64;

```

```

Query Match          13.9%: Score 148.5: DB 1: Length 1966;
Best Local Similarity 21.3%: Pred. No. 0.42:
Matches 54: Conservative 45: Mismatches 97: Indels 57: Gaps 6:

QY 11 LAOEENVL-PREFLNKLNDVNRALQSKDEKRSQYIIDTLRLTEERNATVSLQOA 69
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 892 LVEKTSLFTNNESTKTKTQSDAEERLALFLAQKQKASQKSELNDADNDRIADVGR 951
QY 70 LKAMELCSLTKKKQKYLEEQ-----QDFTKQAOEAGRLSKSKMTM 111
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 952 KKRIAEVVALKKQIQDLEMSLRKAESKQSKDHIIRSLQDEMOQOADAIAKLNEKKHQ 1011
QY 112 EQLLELLQSOLPEVEEMIRDMGVGSAYAEQALAVCVSLKKEVEN----- 156
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1012 EETINKKLEMDLQSE-----DKGNHONKKY-----AKLEQTLDDLEDSLEKKRRADLD 1061
QY 157 KEARKASGEV-----ADLRKDLFFSSKRSKLTQTYVSELDOAKLELSAKNDLQ 203
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1062 KKKRVEEGLKTAQENIDESGRHDLNNILKKSELSHVSRLDEQALVSKLRQIK 1121
QY 204 SADRKMSLKKL 216
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1122 DGQSRISLEEL 1134

RESULT 34
MYSN_DROME
ID_MYSN_DROME STANDARD: PRT: 2017 AA.
AC 099323:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, non-muscle (zipper protein) (Myosin II).
GN GN
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=90349606; PubMed=2117279;
RA Kechum A.S., Stewart C.T., Stewart M., Kiehart D.P.;
RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
RT transcript: conserved sequences in the myosin tail and differential
RT splicing in the 5' untranslated sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
CC -1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
CC CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M35012: AAA28713.1; -.
DR PIR, A36014; A36014.
DR PIR, B36014; B36014.
DR HSSP, P08799; 1MDM.
DR Flybase: FBgn0005634; zip.
DR InterPro: IPR000048; IO.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR001928; Myosin_Tail.
DR InterPro: IPR001609; Myosin_head.

```

DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR Pfam: PF00035; myosin\_head; 1.  
 DR SMART: SM0015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 KW Myosin: Alternative splicing: Coiled coil; Actin-binding;  
 KW ATP-binding; Calmodulin-binding.  
 FT DOMAIN 1 829  
 FT MYOSIN HEAD-LIKE.  
 FT DOMAIN 830 859  
 FT IQ.  
 FT NP\_BIND 886 2017  
 FT COILED COIL (POTENTIAL).  
 FT NP\_BIND 225 232  
 FT ATP.  
 FT DOMAIN 250 260  
 FT 25 KDA/50 KDA JUNCTION.  
 FT DOMAIN 682 694  
 FT 50 KDA/20 KDA JUNCTION.  
 FT DOMAIN 705 727  
 FT ACTIN-BINDING.  
 FT DOMAIN 742 758  
 FT REACTIVE SULFHYDRYL/ACTIN-BINDING.  
 FT DOMAIN 1303 2017  
 FT LIGHT MEROMYOSIN (LMW).  
 FT DOMAIN 1303 1970  
 FT ALPHA-HELICAL TAILPIECE (LMW).  
 FT DOMAIN 1971 2017  
 FT GLOBULAR TAILPIECE.  
 FT VARSPLIC 1 45  
 FT MISSING (IN SHORT ISOFORM).  
 SQ SEQUENCE 2017 AA: 232016 MW: 73E3CB02BA8F2528 CRC64;

Query Match 13.8%; Score 147.5; DB 1; Length 2017;  
 Best Local Similarity 23.7%; Pred. No. 0.49; Mismatches 102; Indels 33; Gaps 7;  
 Matches 57; Conservative 49; Mismatches 102; Indels 33; Gaps 7;

QY 1 RTIINKLEFDLAEEENVLDREFLKNELDNVAQLSQDKER--DSQVI-----ID 50  
 Db 1263 KTVLEKAGCTLEAENAD-----LATELRSVNSSQENDRRKKAQESQIAELQVLALE 1316  
 QY 51 TLRLTLERNATVVSLOQ-----ALCAKEMLCSTLKKQKMYLEQOODETKQAOEENG 103  
 Db 1317 RARSELQER---CTKLOQEAENITNQLQEAELKSAVAKSASNSMSQLTQAEQLEETR 1373  
 QY 104 ---LRSMKMTMEQIELLQSQLPVEVEIMRDMGVGOSAVBOLAVYCVSLKKEYENLKEA 159  
 Db 1374 QKGLSSKLRQTESKEKALQOLEDDDAKKNY---EKRLQEVTTQMOEIKKKAEDDL 1430  
 QY 160 RKASGEVADKLKRLDLEFSSRSKLTQTVYSELQAKLELSAQKDLQADKESIMSLKKLTML 219  
 Db 1431 AKLEEGKKRINKDIEALERQVKELIAQNDRDKSKKIKQSELEDATILEAQRTKVL 1490  
 QY 220 Q 220  
 Db 1491 E 1491

RESULT 35  
 CENE\_HUMAN STANDARD; PRT: 2663 AA.  
 AC Q02224;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Centromeric protein E (CENP-E protein).  
 GN CENPE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:93024922; PubMed:1406971;  
 RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;  
 RT "CENP-E is a putative kinetochore motor that accumulates just before  
 RT mitosis";  
 RL Nature 359:536-539(1992).  
 RN [2]  
 RP CHARACTERIZATION.

RX MEDLINE:95196755; PubMed:7889940;  
 RA Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;  
 RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed  
 RT microtubule motor";  
 RL EMBO J. 14:918-926(1995).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE:98437347; PubMed:9763420;  
 RA Chan G.K.T., Schaar B.T., Yen T.J.;  
 RT "Characterization of the kinetochore binding domain of CENP-E reveals  
 RT interactions with the kinetochore proteins CENP-F and hBUBR1";  
 RL J. Cell Biol. 143:49-63(1998).  
 CC -1- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE  
 CC KINETOCHORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE  
 CC OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT  
 CC AND/OR SPINDLE ELONGATION  
 CC -1- SUBUNIT: INTERACTS WITH CENP-F AND BUBR1 KINASE.  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING  
 CC CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS  
 CC QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.  
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z15005; CAAT8727.1; -.  
 DR EMBL: S28261; S28261.  
 DR HSSP: P17119; 3KAR.  
 DR MIM: 117143; -.  
 DR InterPro: IPR001752; kinesin.  
 DR Pfam: PF00225; kinesin; 1.  
 DR PRINTS: PR00380; KINESINHEAVY.  
 DR SMART: SM00129; KISC; 1.  
 DR PROSITE: PS00411; KINESIN\_MOTOR\_DOMAIN; 1.  
 DR PROSITE: PS50067; KINESIN\_MOTOR\_DOMAIN; 1.  
 KW Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;  
 KW Cell cycle; Centromere.  
 FT DOMAIN 1 335  
 FT KINESIN-MOTOR.  
 FT DOMAIN 336 2471  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 2472 2663  
 FT GLOBULAR (POTENTIAL).  
 FT NP\_BIND 86 93  
 FT ATP (BY SIMILARITY).  
 SQ SEQUENCE 2663 AA: 312087 MW: CEFCl3880C8C8CB8 CRC64;

Query Match 13.8%; Score 147.5; DB 1; Length 2663;  
 Best Local Similarity 25.1%; Pred. No. 0.65;  
 Matches 57; Conservative 48; Mismatches 91; Indels 31; Gaps 9;

QY 13 QEEENVLDREFLKNELDNVAQLSQDKER-RDSQVIITLRLTLERNATVVSLOQALG 71  
 Db 1638 ETOKMEIEHLKEQFOETOKLNTENIETNRLTQI---LHNMLEWR-SVTKRERDLR 1692  
 QY 72 KAEMLCSTLKKQKMYLEQOODET---KQAOEENG---RLRSKMTMEQIELLQSQLP 123  
 Db 1693 SVE---ETLKVERQDLKLENLRETITRDLEKQDELKIVMHMLKEHOETIDKLRGIVSEKTN 1749  
 QY 124 EVEEMIRDMGVGOSAV-----EQLAVYCVSLKKEYENLKEARKASGEVADKL---RK 172  
 Db 1750 EISMNQKDLHSNDAKLAQDLKIEBELRLIAHMLKKEQDETIDKLRGIVSEKTDLSNMQ 1809  
 QY 173 DLFSRSKLTQTVYSELQAKLELSAQKDLQADK---EIMSLKKLT 216  
 Db 1810 DLNSNAKLERQIOELKANEHQTLITLKDVNETQKVSSEMQLKKQI 1856

RESULT 36  
 Y373\_BOVIN STANDARD; PRT: 1453 AA.  
 ID Y373\_BOVIN

AC Q9TU23;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein KIAA0373.  
GN KIAA0373.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovine; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Trachea;  
RA Jovov B., Ripoll P.J., Benos D.J.;  
RT Submitted (MIG-1999) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF176816; AAF00990.1; .  
KW Hypothetical protein; Coiled coil.  
FT DOMAIN 37 1426  
SQ SEQUENCE 1453 AA; 169934 MW; 03CBA02A64CF4139 CRC64;  
  
Query Match 13.8%; Score 147; DB 1; Length 1453;  
Best Local Similarity 23.1%; Pred. No. 0.37; Mismatches 78; Indels 86; Gaps 11;  
Matches 64; Conservative 49; Mismatches 78; Indels 86; Gaps 11;  
  
QY 15 EENVLDREFLKNEI-----DNVRAQLSQRDKERDSQVITDRLDTLEERNATV 64  
DB 571 EENILSRDKVINEIRLRPATAEQEKLAERFSEKREPEPKS-----HHTLKLAHQIA 622  
QY 65 SLQALGKAEMLCTLKQKMYLEQODE---TKQAQEEGRLSKKKTKTEQ----- 113  
DB 623 NMQARLNQKE---EVLKKYQHLLEKAREQREIVKKEHEELHTLRKLELDQADNSLSKFK 679  
QY 114 ---TELLLOS--OLPEVEEMIRDMGVGSAVEQ---LAVYCVSLK----- 150  
DB 680 ETAMDLIKQSTPVPYTKHFRILAMEQTYAEQDDSLSLVTKIKYQSODLEROKETTEL 739  
QY 151 --KEYENLK-----EARKASGEVADKLKDLFSSRSKLQTVYSEL----- 188  
DB 740 KIKFEFENMKLRLOENHADVKKIKAEVED-LRCLLVQSQKESQSLKSELQIQKEANSRAP 798  
QY 189 -----DQAKLELSAQKDLQSADEIKMSLKKLT 217  
DB 799 TTTMRNLVERLKSQALKEKQKALSRALLELRAMET 835  
  
RESULT 37  
RA50\_THEVO STANDARD; PRT; 895 AA.  
AC P58302;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DNA double-strand break repair Rad50 ATPase.  
GN RAD50 OR TVG0235331.  
OS Thermoplasma volcanium.  
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;  
OC Thermoplasma.  
OX NCBI\_TaxID=50339;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GSSI / DSM 4299 / JCM 9571;  
RX MEDLINE=20570466; PubMed=11121031;

RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
RA Kawashima T., Yamamoto Y., Matsumoto K., Yamazaki M., Kanehori K., Kawamoto T.,  
RA Nishishita T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
RT "Archaeal adaptation to higher temperatures revealed by genomic  
RT sequence of Thermoplasma volcanium."  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The  
CC rad50/mre11 complex possesses single-strand endonuclease activity  
CC and ATP-dependent double-strand-specific exonuclease activity.  
CC Rad50 provides an ATP-dependent control of mre11 by unwinding  
CC and/or repositioning DNA ends into the mre11 active site (By  
CC similarity).  
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP000991; BAB59370.1; .  
DR InterPro; IPR003439; ABC\_transportr.  
DR InterPro; IPR000875; Cecropin.  
DR InterPro; IPR001238; RecF.  
DR InterPro; IPR003395; SMC\_N.  
DR Pfam; PF004470; RecF; 1.  
DR Pfam; PF02463; SMC\_N; 1.  
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.  
FT NP\_BIND 30 37  
FT DOMAIN 171 737  
SQ SEQUENCE 895 AA; 103223 MW; 4AA7DA9287E82D3A CRC64;  
  
Query Match 13.7%; Score 146.5; DB 1; Length 895;  
Best Local Similarity 22.8%; Pred. No. 0.24; Mismatches 82; Indels 53; Gaps 8;  
Matches 55; Conservative 51; Mismatches 82; Indels 53; Gaps 8;  
  
QY 14 EENVLDRE-----EFLKNELDNVRAQLSQRDKERDSQVITDRLDTLEERNAT 62  
DB 150 ERKKILDEILNIDRLGAYLLKLEVIDLTANVSDYDLKNELOSKI---NEIDNNKQ 205  
QY 63 VVSLQALGKAEMLCTLKQKMYLEQODETKQAQEEAGLRKMKMTMEQIEL-LQSQ 121  
DB 206 IEELSKRLRLPEIKALEEINIKENKKD--HUNEELRLNQLQLETIKKYEMLAESQ 262  
QY 122 -----LPEVEEMIRDMGVGSAVEQLAVYCVSLKKEYENLKEARKASGEVADK 169  
DB 263 SRKASIEMEVYKLPISIELKRLNNAVYKRNEI-----EYINKKDLGSLSEIES 316  
QY 170 LRKDLF---SSRSKLQTVY-----EIDQAKLELSAQKDLQSADEIKMSL 213  
DB 317 LKSLDSKYDEAHKRLLEDQSFSEFLKRRKREDLKRSLKDEEDNYQSAVNIENIK 376  
QY 214 K 214  
DB 377 K 377  
  
RESULT 38  
MYSS\_RABIT STANDARD; PRT; 1084 AA.  
AC P02562;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Myosin heavy chain, skeletal muscle (Fragments).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;

[1]  
 RA Capony J.-P., Elzinga M.;  
 "The amino acid sequence of A 34,000 dalton fragment from S-2 of  
 myosin.";  
 RL Biophys. J. 33:148A-148A(1981).  
 [2]  
 RX MEDLINE=85131142; PubMed=3972832;  
 RA Lu R.C., Wong A.;  
 "The amino acid sequence and stability predictions of the hinge  
 region in myosin subfragment 2.";  
 RL J. Biol. Chem. 260:3456-3461(1985).  
 [3]  
 RP SEQUENCE OF 409-1084 FROM N.A.  
 RX MEDLINE=87304245; PubMed=3305014;  
 RA Maeda K., Sczakiel G., Wittinghofer A.;  
 "Characterization of cDNA coding for the complete light meromyosin  
 portion of a rabbit fast skeletal muscle myosin heavy chain.";  
 RL Eur. J. Biochem. 167:97-102(1987).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,  
 CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE  
 CC CONSERVED.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X03958; CA29391.1; -;  
 DR PIR; A02985; A02985.  
 DR PIR; A05280; A05280.  
 DR PIR; S00084; S00084.  
 DR HSSP; P02649; IMFN.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KM Ayr-binding; Multigene family.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 258 ALPHA-HELICAL TAILPIECE (S2).  
 FT NON\_CONS 258 259  
 FT DOMAIN <259 1084 RODLIKE TAIL (S2 AND LMW DOMAINS).  
 FT DOMAIN 455 1084 COILED COIL (POTENTIAL).  
 FT VARIANT 405 405 L -> V.  
 FT VARIANT 408 408 V -> L.  
 FT VARIANT 421 421 E -> D.  
 FT VARIANT 423 423 S -> G.  
 FT VARIANT 426 426 K -> R.  
 FT VARIANT 426 426  
 SQ SEQUENCE 1084 AA; 12548 MW; 229CFD69A6E1F70 CRC64;

```

Query Match      13.7%  Score 146;  DB 1;  Length 1084;
Best Local Similarity 24.5%  Pred. No. 0.32;
Matches 47;  Conservative 42;  Mismatches 83;  Indels 20;  Gaps 6;

OY      24  LKNELDNVAALQSKDKERKDSQVITDRLTLEERNATVVSLOQALKAENLCLSTLKQ 83
      ::::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
352  LGEIDINQARKVKLEKESSEIKMIDIDLASNME-----TVSKAKNLEKMCMTLEQ 404

```

QY	84	MYLLEQOODERKQAOEEAGRLRSKMKMTMEQJELLLOLOLPEVEMIMIDMGQSAVBOLA	143
Db	405	LESEKTKKEEHORILINELISAOKARLHT-ESGE--FSROLEKDMVAQSOLSRGGQAFNQ--	459
QY	144	VYCVSLKKEYENLKEARKASGEVADKLRKDFSSRSKLTQTYSELDQAKLELSAOKDLQ	203
Db	460	-QIBGLKQOLEEETKAKSALAHALOSSRDC---DLLREQYEEEOAKAEL--QRAMS	511
QY	204	SADKEIMSILKKK	215
Db	512	KANSEVSOVRTK	523

	RESULT_39			
XCP_C_XENLA	ID	XCPC_XENLA	STANDARD:	PRT: 1290 AA.
AC	PS0532;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Chromosome assembly protein XCAP-C.			
GN	XCAP-C.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95042742; PubMed=7954811;			
RA	Hirano T., Mitchison T.J.,"			
RT	"A heterodimeric coiled-coil protein required for mitotic chromosome			
RL	condensation in vitro.";			
CC	Cell 79:449-458(1994).			
CC	-1- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF			
CC	MITOTIC CHROMOSOMES.			
CC	-1- SUBUNIT: ASSOCIATES WITH XCAP-E PROBABLY AS HETERODIMER.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR. PURIFIED CHROMOSOME ASSEMBLY IN			
CC	MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND			
CC	FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.			
CC	-1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS			
CC	FLANKED BY PUTATIVE GLOBAL REGIONS AT THE N- AND C-TERMINUS.			
CC	-1- SIMILARITY: BELONGS TO THE SMC FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on ways			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: U13673; AAA64679.1; -			
DR	InterPro: IPR003439; ABC_transportr.			
DR	InterPro: IPR003405; SMC_C.			
DR	InterPro: IPR003395; SMC_N.			
DR	Pfam: PF02483; SMC_C; 1.			
KM	Pfam: PF02463; SMC_N; 1.			
DR	Mitosis; ATP-binding; Coiled coil; Nuclear protein.			
FT	NP_BIND 107 114			
FT	DOMAIN 264 594			
FT	DOMAIN 764 1027			
FT	DOMAIN 1094 1129			
FT	DOMAIN 1263 1290			
FT	DOMAIN 60 65			
FT	DOMAIN 747 750			
FT	DOMAIN 841 844			
FT	DOMAIN 1196 1220			
FO	SEQUENCE 1290 AA; 146988 MM; 293124924FFE90F6 CRC64;			
	ALA/ASP-RICH (DA-BOX).			

Query Match	13.7%;	Score 146;	DB 1;	Length 1290;
Best Local Similarity	21.2%;	Pred. No. 0.38;		



```
Db 1257 LSEARGKNEEIQRSLSLTTQKSRLOQTEAGELSRQLEEKESIVSQLSRSKQAFtQOTEEL 1316
Oy 129 IRDMGVGOSAVEOLAVY-----CVSLKKEYENLKEARKASGEVADKLKRDLFSSRSKL 181
Db 1317 KROLEENKAKNALAHALQSSRHDDCLREQYEEOEGKAELORALSKANSEVAQWRTRY 1376
Oy 182 QT---VYSELDOAKLEIKSAQKDQSDAKKEIWSLKKKLTWLO 220
Db 1377 ETDAIQORTEELLEAKKL--AQR-LQDSEEQVEAVNAKCASTLE 1416
```

Search completed: September 4, 2002, 16:17:35  
Job time: 494 sec